HCR-20 Violence Risk Assessment Scheme

Overview and Annotated Bibliography

Kevin S. Douglas, LL.B., Ph.D.
Simon Fraser University Department of Psychology
Mid Sweden University, Applied Criminology

Catherine Shaffer, M.A.
Simon Fraser University Department of Psychology

Adam J. E. Blanchard, M.A.
Simon Fraser University Department of Psychology

Laura S. Guy, Ph.D.
University of Massachusetts Medical School

Kim A. Reeves, Ph.D.
Simon Fraser University Department of Psychology

John Weir, Ph.D.
Southwest Texas Medical School
# Table of Contents

Preface to and Description of this Updated Review and Bibliography ...........................................3

Section 1 | HCR-20⁰⁻³ Studies ................................................................................................................4
   HCR-20⁰⁻³ Research Summaries in this Update (Since January 1, 2014) .........................................5
   Description of the HCR-20⁰⁻³ Violence Risk Assessment Scheme  Last Updated January 1, 2014 ......7
   Table 1. Items in the HCR-20⁰⁻³ Violence Risk Assessment Scheme .............................................10
   Table 2. Summary of HCR-20⁰⁻³ Studies by Sample, Outcome, Gender, and Country ..................12
   Table 3. Summary of Selected HCR-20⁰⁻³ Research: Key Findings .............................................14
   Annotated Bibliography of HCR-20⁰⁻³ Research Projects, Publications, Presentations, and Unpublished Studies ...........................................................................................................17
      Civil Psychiatric Settings ............................................................................................................17
      Forensic Psychiatric Settings .....................................................................................................20
      Correctional Settings ....................................................................................................................35
      Mixed Settings ............................................................................................................................36
   HCR-20⁰⁻³ References ....................................................................................................................39
   Scholarly Works that were Summarized in Section 1 (followed by the page number on which they appear) ..................................................................................................................39
   Other Scholarly Works Cited in Section 1 ....................................................................................41
   Awaiting Copies of the Following Presentations ........................................................................42
   HCR-20⁰⁻³ Related Disseminations and Academic Commentary ..................................................43
   HCR-20⁰⁻³ Professional Manuals ....................................................................................................43
   HCR-20⁰⁻³ Risk Factor Literature Review and Rationale ...............................................................43
   HCR-20⁰⁻³ Translations ....................................................................................................................43
   Related HCR-20⁰⁻³ Chapters and Articles by HCR-20 Authors .......................................................43
   Published HCR-20⁰⁻³ Peer Academic Commentary ....................................................................43

Section 2 | HCR-20¹⁻² Studies ..............................................................................................................45
   Description of the HCR-20¹⁻² Violence Risk Assessment Scheme .............................................46
   Table 4. Items in the HCR-20¹⁻² Violence Risk Assessment Scheme .............................................47
   Receiver Operating Characteristic (ROC) Analyses .................................................................48
   Table 5. Summary of HCR-20¹⁻² Studies by Sample, Outcome, Gender, and Country .............50
   Table 6. A Summary of Selected HCR-20¹⁻² Research: Key Findings ........................................53
   Annotated Bibliography of HCR-20¹⁻² Research Projects, Publications, Presentations, and Unpublished Studies ..............................................................................................................71
      Civil Psychiatric Settings ............................................................................................................71
      Forensic Psychiatric Settings .....................................................................................................88
      Correctional Settings ..................................................................................................................153
      Mixed Settings ............................................................................................................................177
      Juvenile Settings ........................................................................................................................195
      Case Law Reviews .....................................................................................................................197
   HCR-20¹⁻² References ....................................................................................................................209
   Scholarly Works that were Summarized in Section 2 (followed by the page number on which they appear) .................................................................209
   Other Scholarly Works Cited in Section 2 ....................................................................................223
Awaiting Copies of the Following Presentations ............................................. 226
Awaiting Copies of the Following Publications ............................................. 227
HCR-20^V1/V2 Related Disseminations and Academic Commentary ............... 228
HCR-20^V1/V2 Professional Manuals ............................................................. 228
HCR-20^V1/V2 Risk Factor Literature Review and Rationale ........................... 228
HCR-20^V1/V2 Book and Test Reviews ......................................................... 228
Meta-Analytic Reviews in which the HCR-20^V1/V2 is Included ..................... 228
HCR-20^V1/V2 Translations ........................................................................... 229
Related Chapters and Articles by HCR-20^V1/V2 Authors ............................... 229
Select HCR-20^V1/V2 Published Peer Academic Commentary ........................ 230
Related HCR-20^V1/V2 Conference Presentations ......................................... 232
Related Professional Manuals ....................................................................... 234
An Invitation and Request ............................................................................. 235
Contact Information ....................................................................................... 236
Preface to and Description of this Updated Review and Bibliography

Last Updated January 1, 2017

Since the publication of Version 1 in 1995 and Version 2 in 1997, the HCR-20 has become one of the world’s most widely used and best validated violence risk assessment instruments. It has been translated into 20 languages and adopted or evaluated in more than 35 countries. Version 3, developed over 5 years on the basis of extensive clinical beta testing and empirical evaluation, was released in 2013.

This extensive HCR-20 Review and Annotated Bibliography identifies significant research carried out with the three versions of the HCR-20. This review has been maintained and updated since the mid-1990s and since then has grown to include 22 disseminations with the V3 (with many more underway) and 233 disseminations with earlier versions (with updates of recent years coming). To further facilitate its use, the following features have been added to the review:

1. Separate sections for HCR-20\(^{V3}\) and HCR-20\(^{V1/V2}\) studies, including separate summary tables and references
2. A description of the HCR-20\(^{V3}\) and a table of revisions to the items
3. New sections on HCR-20\(^{V1/V2}\) case law review and non-empirical studies
4. Combined civil and general psychiatric sections

Stay tuned for updates!
Section 1 | HCR-20$^\text{V3}$ Studies
**HCR-20 V3 Research Summaries in this Update**  
*(Since January 1, 2014)*


Description of the HCR-20\textsuperscript{V3} Violence Risk Assessment Scheme

Last Updated January 1, 2014

Version 3 of the HCR-20 (Douglas, Hart, Webster, & Belfrage, 2013) embodies and exemplifies the Structured Professional Judgment (SPJ) approach to violence risk assessment and management. In addition to this fact, we were guided by several principles in its revision – (a) continuity of concept (V3 has a similar “look and feel” compared to V2); (b) clinical and practical utility (the ultimate purpose of V3 is to help practitioners evaluate and manage risk of individual patients, clients, or offenders); (c) enhanced clarity (we clarified any areas from Version 2 that had been pointed out to use as being unclear); (d) legal and ethical acceptability (we exclude objectionable risk factors; the process of risk assessment outlined in V3 allows for clinical and judicial review of the process of risk assessment); (e) empirical defensibility (we conducted or asked others to conduct beta-testing and empirical evaluation of V3 prior to publishing it).

Although V2 performs well and has been widely adopted, based on the guiding principles outlined above we believed we could improve it. One of the major goals was to ensure that it fully exemplified contemporary SPJ scholarship. A lot of thinking and work on risk assessment and management has transpired since we published V2 in 1997. V3 now embodies this work. We engaged in extensive consultation, examination of the literatures on violence and risk assessment generally and the HCR-20 more specifically, in-depth beta-testing and feedback, and considerable empirical evaluation in the development of V3. A number of our colleagues had shared their HCR-20 V2 data with us so we could evaluate its performance in large-scale samples (i.e., 5000+ cases). This helped to guide us in terms of whether certain items should be revised or added. Further, starting with Chris Webster’s work in the early 1990s, the HCR-20 approach has always sought to incorporate the ideas of working clinicians and practitioners – people who understand what works on the ground, and what is and is not helpful to practice. We retained this “grass roots” element in the development of Version 3.

Although we did make a number of meaningful changes, we believe and have received feedback that those who are familiar with the SPJ approach more broadly or with HCR-20 V2 more specifically can shift to using V3 quite comfortably. A number of the additions and changes reflect what well-trained clinicians were doing in practice anyways.

HCR-20 V3 is described in detail in several publications (and the manual, of course!), but we highlight some of the major changes here. First, the risk factors are presented in Table 1. As is clear, one of the changes was the addition of sub-items for broader or more complicated risk factors. These help evaluators and decision-makers specify and conceptualize the nature of risk more precisely. We did also make some changes to several items, including dropping one or two, and adding one or two others. We also now ask that evaluators rate not only the presence of risk factors, but also their relevance to the
violence of the individual being evaluated. This step helps evaluators in the next (new) steps of V3 – case formulation and scenario planning. In these steps, evaluators are provided with guidance in terms of developing formulations that explain why a person has acted violently, and what they are concerned a person might do in the future. To further aid in formulation, we have included sets of “indicators” for each risk factor. These indicators are there for guidance, and provide specific ways in which risk factors might manifest at the individual level. There is also now more explicit attention paid to generating risk management and risk reduction plans that link to relevant risk factors, formulations, and scenario plans. Another change was the removal of the requirement to use either the Psychopathy Checklist Revised or Screening Version. Evaluators can still use these if they would like to, but they are not required to. Finally, we have included the opportunity for evaluators to make not only overall summary risk ratings, but summary risk ratings of serious violence and of imminent violence.

We have produced three rating sheet options. The use of any of these is discretionary, not required. There remains a simple 1-page rating sheet (presence and relevance of risk factors; summary risk ratings), and a 2-page rating sheet (presence and relevance of risk factors and sub-items; summary risk ratings). These are available for free download from HCR-20.com. We have also produced an extended worksheet that includes the multistep HCR-20 V3 procedure (i.e., including sections for formulation, scenario planning, risk management).

We are currently finalizing several papers for the HCR-20 White Paper Series, of which this Annotated Bibliography is White Paper #1. In addition, we have made available a 300-page violence literature review prepared by Guy and Wilson (2007) that we used as part of our start to the revision process (White Paper #2). We are finalizing White Paper #3 (Guy et al., in prep), which is an item-by-item literature review for each HCR-20 V3 risk factor that summarizes the literature in its support, and the mechanisms by which it might lead to violence. In addition, we will be revising the HCR-20 Violence Risk Management Companion Guide (Douglas et al., 2001), which is an edited compilation of chapters that align with HCR-20 risk factors and that provide suggestions for risk reduction and management. Finally, we have recruited a number of colleagues to put together a special issue on HCR-20 V3 articles. This is currently under review and we will provide summaries of the articles in this document once it is published.

Proper Citations


Table 1. Items in the HCR-20 V3 Violence Risk Assessment Scheme

<table>
<thead>
<tr>
<th>Sub-Scales</th>
<th>Items</th>
<th>Sub-Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Historical Scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td>History of Problems With Violence</td>
<td>As a Child (12 and under)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>As an Adolescent (13 – 17)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>As an Adult (18 and over)</td>
</tr>
<tr>
<td>H2</td>
<td>History of Problems With Other Antisocial Behavior</td>
<td>As a Child (12 and under)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>As an Adolescent (13 – 17)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>As an Adult (18 and over)</td>
</tr>
<tr>
<td>H3</td>
<td>History of Problems With Relationships</td>
<td>Intimate Relationships</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-Intimate Relationships</td>
</tr>
<tr>
<td>H4</td>
<td>History of Problems With Employment</td>
<td>--</td>
</tr>
<tr>
<td>H5</td>
<td>History of Problems With Substance Use</td>
<td>--</td>
</tr>
<tr>
<td>H6</td>
<td>History of Problems With Major Mental Disorder</td>
<td>Psychotic Disorders</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Major Mood Disorders</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other Major Mental Disorders</td>
</tr>
<tr>
<td>H7</td>
<td>History of Problems with Personality Disorder</td>
<td>Antisocial, Psychopathic and Dissocial</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other</td>
</tr>
<tr>
<td>H8</td>
<td>History of Problems With Traumatic Experiences</td>
<td>Victimization/Trauma</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adverse Childrearing Experiences</td>
</tr>
<tr>
<td>H9</td>
<td>History of Problems With Violent Attitudes</td>
<td>--</td>
</tr>
<tr>
<td>H10</td>
<td>History of Problems With Treatment or Supervision Response</td>
<td>--</td>
</tr>
</tbody>
</table>
### Sub-Scales  | Items                                      | Sub-Items                  
---            |-------------------------------------------|-----------------------------

#### Clinical Scale

| C1            | Recent Problems With Insight              | Mental Disorder             |
|               |                                           | Violence Risk               |
|               |                                           | Need for Treatment          |
| C2            | Recent Problems With Violent Ideation or Intent | --                         |
| C3            | Recent Problems With Symptoms of Major Mental Disorder | Psychotic Disorders          |
|               |                                           | Major Mood Disorders        |
|               |                                           | Other Major Mental Disorders|
| C4            | Recent Problems with Instability          | Affective                   |
|               |                                           | Behavioral                  |
|               |                                           | Cognitive                   |
| C5            | Recent Problems with Treatment or Supervision Response | Compliance                  |
|               |                                           | Responsiveness               |

#### Risk Management Scale

| R1            | Future Problems With Professional Services and Plans | --                          |
| R2            | Future Problems With Living Situation             | --                          |
| R3            | Future Problems With Personal Support             | --                          |
| R4            | Future Problems With Treatment or Supervision Response | Compliance                  |
|               |                                                   | Responsiveness               |
| R5            | Future Problems With Stress or Coping             | --                          |

*Note. Adapted from Douglas, Hart, Webster & Belfrage (2013).*
Table 2. Summary of HCR-20\textsuperscript{V3} Studies by Sample, Outcome, Gender, and Country

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Civil Psychiatric</th>
<th>Forensic Psychiatric</th>
<th>Correctional</th>
<th>Mixed</th>
<th>Juvenile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Entries</td>
<td>Total N</td>
<td>Entries</td>
<td>Total N</td>
<td>Entries</td>
<td>Total N</td>
</tr>
<tr>
<td>By Outcome\textsuperscript{1}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Institution</td>
<td>9</td>
<td>687</td>
<td>1</td>
<td>156</td>
<td>8</td>
<td>531</td>
</tr>
<tr>
<td>Community</td>
<td>4</td>
<td>688</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>582</td>
</tr>
<tr>
<td>Both</td>
<td>1</td>
<td>99</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>99</td>
</tr>
<tr>
<td>By Sample Gender\textsuperscript{2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>13</td>
<td>1,348</td>
<td>1</td>
<td>156</td>
<td>11</td>
<td>1,192</td>
</tr>
<tr>
<td>Male</td>
<td>2</td>
<td>104</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Female</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

1 The number of entries and total $N$ do not always add up to match the totals row due to missing information from some studies.
2 The number of entries and total $N$ do not always add up to match the totals row due to missing information from some studies. Also, it is acknowledged that many of the studies that utilized mixed gender samples contained predominantly males.
<table>
<thead>
<tr>
<th>Country</th>
<th>Total Entries</th>
<th>Total N</th>
<th>Civil Psychiatric Entries</th>
<th>Total N</th>
<th>Forensic Psychiatric Entries</th>
<th>Total N</th>
<th>Correctional Entries</th>
<th>Total N</th>
<th>Mixed Entries</th>
<th>Total N</th>
<th>Juvenile Entries</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>6</td>
<td>439</td>
<td></td>
<td></td>
<td>5</td>
<td>333</td>
<td>1</td>
<td>106</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>3</td>
<td>345</td>
<td>1</td>
<td>156</td>
<td>1</td>
<td>105</td>
<td>1</td>
<td>84</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>2</td>
<td>283</td>
<td></td>
<td></td>
<td>2</td>
<td>283</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>1</td>
<td>409</td>
<td></td>
<td></td>
<td>1</td>
<td>409</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>1</td>
<td>32</td>
<td></td>
<td></td>
<td>1</td>
<td>32</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>1</td>
<td>30</td>
<td></td>
<td></td>
<td>1</td>
<td>30</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>1</td>
<td>20</td>
<td></td>
<td></td>
<td>1</td>
<td>20</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>International</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Portugal</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serbia</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Zealand</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Romania</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scotland</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greece</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The number of entries and total N do not always add up to match the totals row due to missing information from some studies.
Table 3. Summary of Selected HCR-20\textsuperscript{V3} Research: Key Findings

<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>Means (SD)</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>HCR Total</td>
<td>H Scale</td>
<td>C Scale</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Civil Psychiatric</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Howe et al. (2016) &amp; 1</td>
<td>64</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wijetunga et al. (2016)</td>
<td>156</td>
<td>24.60\textsuperscript{A}</td>
<td>11.76\textsuperscript{A}</td>
<td>7.00\textsuperscript{A}</td>
</tr>
<tr>
<td>Forensic Psychiatric</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bjorkly, Eidhammer &amp; Selmer (2014)</td>
<td>20</td>
<td>25.50 (3.61)</td>
<td>13.45 (3.28)</td>
<td>6.35 (1.23)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cook et al. (2014)</td>
<td>39</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Vries Robbé &amp; de Vogel (2015)</td>
<td>197</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Vogel &amp; van de Broek (2014)</td>
<td>86</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doyle et al. (2014)</td>
<td>387</td>
<td>22.48 (6.35)</td>
<td>13.04 (3.68)</td>
<td>4.84 (2.66)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Douglas &amp; Belfrage (2014)</td>
<td>32</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Continues on Next Page with More Forensic Samples
<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>HCR Total</th>
<th>H Scale</th>
<th>C Scale</th>
<th>R Scale</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forensic Psychiatric (Continued)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Green et al. (2015)</td>
<td>97</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUCs = .73, .74, .61</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(H, C, R)</td>
<td></td>
</tr>
<tr>
<td>Griswold et al. (2016)</td>
<td>28</td>
<td>27.32 (6.53)</td>
<td>15.11 (3.05)</td>
<td>5.57 (2.18)</td>
<td>6.64 (2.93)</td>
<td>AUCs = .81, .72, .83, .75</td>
<td>ICC (SPJ) = .91</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Total, H, C, R)</td>
<td></td>
</tr>
<tr>
<td>Hogan &amp; Olver (2015)</td>
<td>99</td>
<td>24.69 (7.27)</td>
<td>12.58 (4.08)</td>
<td>5.12 (2.94)</td>
<td>6.99 (1.93)</td>
<td>AUCs = .76, .64, .76, .76</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Total, H, C, R)</td>
<td></td>
</tr>
<tr>
<td>Kötter et al. (2014)</td>
<td>30</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
<td>ICC (SPJ) = .86²</td>
</tr>
<tr>
<td>Penney, Marshall, &amp; Simpson</td>
<td>87</td>
<td>19.78 (4.40)</td>
<td>12.52 (2.55)</td>
<td>3.33 (1.99)</td>
<td>3.92 (1.40)</td>
<td>Odds ratios = 1.55, 0.78, 1.68</td>
<td>ICC range (12 cases) = .84-85</td>
</tr>
<tr>
<td>(2016)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(H, C, R)</td>
<td></td>
</tr>
<tr>
<td>Schneider et al. (2015)</td>
<td>106</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUCs = .75, .67, .78, .68</td>
<td>ICCs = .82, .86, .84</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Total, H, C, R)</td>
<td>(H, C, R)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Odds ratios = 1.12, 1.90, 0.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(H, C, R)</td>
<td></td>
</tr>
<tr>
<td>Strub &amp; Douglas (2009)</td>
<td>80</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>r = .60 with HCR-20 V2 (H)</td>
<td>ICC₁ = .75, ICC₂ = .85 (H)</td>
</tr>
<tr>
<td><strong>Correctional Samples</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smith et al. (2014)</td>
<td>84</td>
<td>--</td>
<td>13.92 (4.18)</td>
<td>5.43 (2.23)</td>
<td>6.92 (2.32)</td>
<td>AUCs = .75, .74, .61</td>
<td>ICCs (15 Cases) = .92, .62, 88</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(H, C, R)</td>
<td></td>
</tr>
</tbody>
</table>

Continues on Next Page with Mixed Samples
<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>Means (SD)</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>HCR Total</td>
<td>H Scale</td>
<td>C Scale</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed Samples</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blanchard &amp; Douglas (2011) (^\psi), 14</td>
<td>43</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Strub, Douglas, &amp; Nicholls (2014)</td>
<td>106</td>
<td>22.9 (7.8)</td>
<td>12.8 (4.3)</td>
<td>4.9 (2.5)</td>
</tr>
</tbody>
</table>

**Note 1.** This table does not contain all studies reported in the annotated bibliography. Some studies supplemental to main studies were not included. Other studies were excluded from the table if they addressed issues other than the relationship between the HCR-20 V1/V2 and violence. The method and results of the studies in this table are described in more detail in the annotated bibliography that follows.

**Note 2.** IRR = interrater reliability; SPJ = structured professional judgment of low, moderate, or high risk; Superscript “A” denotes analyses for men only; Superscript “B” denotes analyses for women only. “\(^\psi\)” denotes that the sample is a sub-sample of another study, and hence the reporting of Ms and SDs is omitted.

2. The authors also report ICC values for individual HCR-20 V3 items.
3. AUC values reported are for any violence. The authors also provide AUC values for physical violence.
4. AUC values are also reported separately for males and females and separately for patients that intramural or on supervised leave from the hospital and patients that are transmural or on unsupervised leave from the hospital.
5. SPJ ratings were given on 3-point and 5-point scales. AUC values reported are for violent recidivism at 3-years. The authors also provide AUC values for violent recidivism at 1- and 2-years.
6. AUC values reported are for violence at 12-months. The authors also provide AUC values for violence at 6-months.
7. The ICC value provided is for R Out rating. The authors also provide the ICC value for R In rating.
8. The authors also report AUC values for the association between change scores on the C, R, and combined C and R scales and violent reoffending.
10. Means reported are for the baseline assessment. The authors also provide means for the discharge assessment.
11. Means reported are for the baseline assessment. The authors also provide means of C and R subscales at 1- and 6-month assessments.
12. Cronbach’s alpha values provided are for the baseline assessment. The authors also provide Cronbach’s alpha values at the 6-month assessment.
13. Mean and ICC value provided is for R Out rating. The authors also provide the mean and ICC value for R In rating.
15. Mean value provided is for R Out rating. The authors also provide the mean value for the R In rating.
16. Mean and AUC values are also reported separately for for psychiatric and offender sub-samples. AUCs reported are for violence at 6-8 months. AUC values are also reported at 4-6 weeks.
Annotated Bibliography of HCR-20 V3 Research Projects, Publications, Presentations, and Unpublished Studies

Civil Psychiatric Settings

Project and Scholarly Works


Summary

The current study was conducted to (a) evaluate the interrater reliability of the HCR-20 Version 3 items and summary risk rating (SRRs), (b) to examine associations between corresponding subscales and SRRs on the HCR-20 Version 3 and Version 2, and (c) to explore how HCR-20 Version 3 presence and relevance ratings were associated with summary risk rating (SRRs). Participants included 64 civil psychiatric patients in the United States. Age ranged from 18 to 70 years with an average of 37.97 years (SD = 13.15). A majority of the sample was male (73%, n = 46), African-American (69%, n = 44), and had a principal diagnosis of schizophrenia or schizoaffective disorder (86%, n = 59).

Violence risk ratings on the HCR-20 Version 2 and Version 3 were completed, on average, about 4 weeks after the patient was admitted to the hospital by trained research assistants. All raters rated both Version 2 and Version 3 making independent ratings for each scale and scoring both instruments based on information contained in the patient’s medical records and, when possible, brief interviews with the patient’s treatment team. Of the 64 participants included in this study, 35 cases were also rated independently by two raters. In order to compare ratings on the two versions, presence and relevance ratings were converted to numerical ratings (0 = not present/low relevance, 1 = possibly or partially present/moderate relevance, and 2 = present/high relevance).

With respect to the HCR-20 Version 2 presence ratings, mean scores were as follows: Total (M = 26.90, SD = 5.90), Historical subscale (M = 11.71, SD = 3.32), Clinical subscale (M = 7.06, SD = 1.89), and Risk Management subscale (M = 8.07, SD = 1.85). With respect to the HCR-20 Version 3 presence ratings, mean scores were as follows: Total (M = 28.26, SD = 5.55), Historical subscale (M = 12.83, SD = 3.12), Clinical subscale (M = 7.46, SD = 1.92), and Risk Management subscale (M = 7.96, SD = 1.85).

First, the authors calculated intraclass correlation coefficients (ICCs; two-way, absolute agreement) to evaluate interrater reliability. On Version 2, good to excellent interrater reliability was observed for the HCR-20 Total (ICC = .80), H subscale (ICC = .84), C subscale (ICC = .70), and R subscale (ICC = .60), and Final Judgement (ICC = .66). On Version 3, good interrater reliability was observed for HCR-20 Total Presence (ICC = .64) and H subscale Presence (ICC = .71). Moderate reliability was observed C subscale Presence (ICC = .55) and R subscale Presence (ICC = .48). With regard to the three summary risk ratings (SRRs), the Imminent Violence rating had excellent interrater reliability (ICC = .77), whereas reliability for both the Future Violence/Case Prioritization and Serious Physical Harm ratings was somewhat lower (ICC = .51 and = .57, respectively). ICC values for individual V2 and V3 items were also presented.

Second, the authors examined the association between Version 2 and Version 3 using Pearson’s correlation coefficients. First, the authors compared Version 2 and Version 3 ratings when two independent raters rated both instruments. Correlations between Total scores on Version 2 and Version 3, (r = .62, p < .001), as well as H (r = .68, p < .001), C (r = .48, p < .01) and R subscales was significant (r = .46, p < .01, respectively). In addition, Version 2 Final Judgements were significantly correlated with all three SRRs on Version 3 including Future Violence/Case Prioritization (r = .40, p < .05), Serious Physical Harm (r = .39, p < .05), and Imminent Violence (r = .58, p < .001). Second, the authors compared the association between Version 2 and Version 3 rated by the same rater. Correlations between Total scores (r = .90, p < .001), as well as correlations between the H (r = .84, p < .001), C (r = .80, p < .001), and R subscales (r = .85, p < .001) were
significant. Version 2’s Final Judgement was also significantly correlated with that rater’s Version 3’s Future Violence/Case Prioritization ($r = .67, p < .001$), Serious Physical Harm ($r = .71, p < .001$), and Imminent Violence ratings ($r = .75, p < .001$).

Last, correlational analyses were used to estimate the associations between HCR-20 Version 3 presence and relevance ratings and the three SRRs. In addition, to explore how presence and relevance ratings may interact, these ratings were multiplied together to create an interaction term. Correlations between presence ratings and case prioritization ratings ranged from .14 (H5: Substance Use) to .53 (H1: Violence). Correlations between relevance ratings and case prioritization ratings ranged from .01 (H9: Violent Attitudes) and .38 (H1: Violence). Correlations between the Presence × Relevance interaction and case prioritization ratings ranged from .15 (R1: Professional Services) to .61 (H1: Violence). A similar pattern of correlations was observed between item presence, relevance, and interaction ratings and serious physical harm and imminent violence.


Summary

The assessment of substance use as a risk factor for aggressive and violent behavior has been well demonstrated in the literature. However, the influence of specific substances on violence risk decision-making is unclear. Using data collected from 77 civil psychiatric inpatients who had been diagnosed with a substance use disorder or self-reported substance use, the authors examined the association between specific abused substances and severity of abuse with the HCR-20 V3. Of the sample, 17 patients (22.1%) were diagnosed with alcohol use disorder, 35 (45.5%) used marijuana, and 26 (33.7%) used hard drugs (e.g., cocaine, heroin, PCP).

First, the authors examined differences in descriptive characteristics across the three groups. Drug users did not significantly differ with respect to gender, ethnicity, or education ($p > .05$). However, cannabis users ($M = 30.40, SD = 10.80$) were significantly younger (all $p < .05$) than alcohol ($M = 42.3, SD = 14.2$) and other drug users ($M = 39.7, SD = 13.8$). In addition, cannabis users (88%) were more likely to be single compared to alcohol users (56%, $p < .05$). The difference in single marital status between other drug users and alcohol or cannabis users was not significant ($p > .05$).

Second, the authors examined differences across the three groups in rates of violent behavior, threats, and thoughts in the prior three month period. Although alcohol users had lower rates of violent behavior (18% vs. 32% and 19%), violent threats (24% vs. 46% and 50%), and violent thoughts (18% vs. 32% and 23%) compared to cannabis and other drug users, this difference was not significant ($p > .05$).

Third, the authors examined whether scores on H5 Presence, H5 Presence*Relevance, and Total HCR-20 Total*Relevance ratings, as well as Summary Risk Ratings (i.e., Future Violence/Case Prioritization; Serious Physical Harm, and Imminent Violence) varied as a function of drug use. Compared to cannabis users ($M = 1.3, SD = 0.8$), other drug users had significantly higher scores on H5 Presence Ratings ($M = 1.8, SD = 0.40, p < .05$). Similarly, other drug users had significantly higher scores on H5 Presence*Relevance Ratings ($M = 4.1, SD = 1.7$) compared to cannabis users ($M = 2.6, SD = 2.4, p < .05$). No other differences were significant ($p > .05$).

Fourth, the authors examined whether differences in rates of violent behavior and HCR-20 V3 scores varied as a function of self-reported substance use or substance use diagnosis. Violent behavior, threats, and thoughts did not significantly vary as a function of diagnostic status ($p > .05$). However, compared to participants diagnosed with a substance use disorder, participants who self-reported substance use had lower scores on H5 Presence ($M = 1.2, SD = 0.8$ vs. $M = 1.9, SD = 0.8, p < .001$), and H5 Presence*Relevance ratings ($M = 2.6, SD = 2.3$ vs. $M = 4.0, SD = 1.9, p < .01$). The authors concluded that the use of specific substances influence risk assessment ratings on the HCR-20 V3.


Summary

Research has generally supported the predictive accuracy and clinical utility of the structured professional judgment (SPJ) approach to violence risk assessment. However, most of this research has focused on men, and the few studies that have examined the SPJ approach in women have yielded mixed results. The current study was conducted to examine gender differences in ratings on the HCR-20 V3 in a male and female civil psychiatric sample. Participants were comprised of 156 adult psychiatric inpatients (42.3%, $n = 66$ female).
The HCR-20 was coded based on available interview data and medical records. To examine interrater reliably, a subset of 35 randomly selected cases was coded by a second rater. ICCs were .85, .73, and .63 for Future Violence/Case Prioritization, Serious Physical Harm, and Imminent Violence summary risk ratings (SRRs). ICCs for HCR-20 total and subscale scores were not reported. Among men, mean scores were as follows: Total Presence (M = 24.60), H Presence (M =11.76), C Presence (M = 7.00), and R Presence (M = 5.84). Among women, mean scores were as follows: Total Presence (M = 21.35), H Presence (M = 9.98), C Presence (M = 6.38), and R Presence (M = 4.98). Standard deviations were not reported.

First, the authors examined gender differences in HCR-20 V3 SRRs. Compared to men, women were significantly more likely to be rated as low risk on all three HCR-20 V3 SRRs (p < .05).

Second, the authors examined gender differences in HCR-20 V3 presence total and scale scores. HCR-20 V3 presence total and scale scores did not significantly differ between men and women (p > .05). However, when HCR-20 V3 scores were calculated by summing the product of item presence and relevance ratings, women’s total and subscale scores were significantly lower than men’s scores (p < .05).

Third, the authors examined gender differences in HCR-20 V3 item scores. Compared to women, men scored higher on the following items: H1 (Violence), H2 (Antisocial Behavior), H9 (Violent Attitudes), C2 (Violent Ideation/Indent) and R1 (Professional Services) (p < .05). Compared to men, women scored higher on H8 (Trauma) (p < .05). No other differences were significant (p > .05).

Last, the authors examined gender differences in the association between HCR-20 V3 scores and SRRs. For women, fewer total, subscale, and items scores significantly correlated with the SRRs than for men, and many of these correlations were significantly weaker. When relevance was considered, the number and strength of significant item-SRR correlations increased less for women than for men. Moreover, the items most strongly associated with SRRs differed between women and men. Taken together, these results suggest that clinicians might evaluate the salience of certain violence risk factors based on gender, and that item ratings might influence their final risk judgments for women less than they do for men. However, such conclusions must await replication before being accepted with a high degree of confidence.

End of Civil Psychiatric Settings
Summary

The current study compared the clinical application of the HCR-20 V3 with V2. Ratings were carried out in a forensic medium security unit in Norway. Two psychiatric nurses compared the V2 and V3 by assessing 20 male forensic psychiatric patients. The raters independently assessed half of the patients each. First, they made a complete assessment of the patients with the V2. After that the same procedure was followed with V3 for the same patients. Assessment data was gathered from patient files, observations, and consulting colleagues. Because items in V2 are scored 0, 1, 2, and items in V3 are coded y (yes), p (possibly), n (no), to obtain data for statistical analysis a common scale of 0, 1, and 2 was chosen in consultation with the HCR-20 V3 lead author to transform V3 ratings into scores.

With respect to the V2, mean scores were as follows: Historical subscale (M = 12.10, SD = 3.92), Clinical subscale (M = 4.90, SD = 1.41), Risk Management subscale (M = 5.80, SD = 1.85), and Total score (M = 22.80, SD = 4.63). With respect to V3 presence ratings, mean scores were as follows: Historical subscale (M = 13.45, SD = 6.35), Clinical subscale (M = 6.35, SD = 1.23), Risk Management subscale (M = 5.70, SD = 2.00), and Total score (M = 22.50, SD = 3.61). With respect to V3 relevance ratings, mean scores were as follows: Historical subscale (M = 13.10, SD = 3.39), Clinical subscale (M = 7.25, SD = 1.41), Risk Management subscale (M = 7.20, SD = 2.04), and Total score (M = 27.55, SD = 3.93).

There was a significant correlation between V2 and V3 Total scores for both presence (r = .58, p < .01) and relevance ratings (r = .55, p < .01). However, paired samples t-tests indicated significant score differences between the two versions. With respect to presence ratings, score differences were observed for the Historical subscale (Cohen’s d = 0.37, p < .05) the Clinical subscale (Cohen’s d = 1.09, p < .01), and Total score (Cohen’s d = 0.65, p < .01), but not for the Risk Management subscale (Cohen’s d = 0.05, p > .05). With respect to relevance ratings, significant score differences were observed for the Historical subscale (Cohen’s d = 0.27, p < .05), Clinical subscale (Cohen’s d = 1.67, p < .001), Risk Management subscale (Cohen’s d = 0.88, p < .05) and Total score (Cohen’s d = 1.10, p < .001).

Findings from an interrater reliability test of V2 and V3 were also reported. There was moderate (Clinical subscale) to good (Historical subscale, Risk Management subscale, and Total score) estimates of internal consistency between the two versions of the HCR-20. With respect to presence ratings, IRR values were .85, .59, .81, and .84 for Historical, Social Contextual, Risk Management and Total scores of V2 and V3, respectively. With respect to relevance ratings ICC values were .85, .58, .79, and .84 for Historical, Clinical, Risk Management and Total scores of V2 and V3, respectively. The fact that scores on the C items yielded lower internal consistency when comparing the two versions was taken to indicate that the most substantial difference in V3 pertains to clinical items. Using a case illustration, HCR-20 evaluations concerning risk factors and judgment procedures using V2 and V3 were also reported.

The authors concluded that the two versions reflect common underlying dimensions; however there were still differences between V2 and V3 ratings for the same patients. The authors found that overall, compared to V2, that the V3 contributed to more systematic and detailed violence risk assessment, with enhanced opportunity to conduct accurate, individual violence risk assessment. The introduction of the new risk assessment category (“Relevance”) to emphasize individual risk factors was of major asset of V3. The authors also comment on the coding and risk formulation of V3.

See Also


Project and Scholarly Works

Summary

Prior research on the predictive accuracy of risk assessment instruments has typically utilized two time points; risk measured at Time 1 and violence measured at Time 2. This method assures temporal ordering of risk factors and violent behavior, but it does not take into consideration that some risk factors change and may no longer be relevant when violence is measured. With these limitations in mind, the current prospective study was conducted to examine the predictive efficacy of HCR-20 V3 and Structured Assessment of Protective Factors for Violence Risk (SAPROF) across multiple time points. Study data were derived from 409 male and female patients discharged from medium secure services in the UK. HCR-20 V3 and SAPROF assessments were completed on the basis of file information and collateral interviews. With the exception of the HCR-20 V3 Historical subscale, which was coded at the baseline assessment only, HCR-20 and SAPROF assessments were completed at baseline, prior to or shortly after discharge, and at 6 and 12-months post-discharge. Violence at 6- and 12-month post-discharge was measured using the MacArthur Community Violence Instrument, using information collected through participant self-report and police records. Inter-rater reliability of the HCR-20 V3 and SAPROF was calculated between four raters on a subsample of 20 cases. IRR values were .92 and .98 for HCR-20 V3 and SAPROF total scores, respectively.

First, the authors examined predictive accuracy of HCR-20 V3 Historical subscale for violence over the follow-up period using Receiver Operating Characteristics (ROC) analysis and mixed effects models. Using ROC analyses, 4 of the 10 Historical items and Historical subscale scores were significantly associated with violence ($p < .05$). A similar pattern of findings was obtained when results were examined using mixed effects models. However, when age, gender, ethnicity, and baseline diagnosis were controlled for in these analyses, only 2 of the 10 items and Historical subscale scores were significantly associated with violence ($p < .05$).

Second, the authors examined predictive accuracy of the HCR-20 Clinical and Risk Management subscales, and the SAPROF. Two separate models were tested using ROC-AUC analyses. First, the authors examined the association between risk/protective factors measured in the past six months to predict violence in the subsequent six months (lagged model) and second, examined the co-occurrence between risk/protective scores and violent behavior within the same six month time window (temporal proximity model). With respect to the lagged model, 6 of the 10 items HCR-20 C and R items, and subscale scores were significantly associated with violence ($p < .05$). In addition, 13 of the 17 SAPROF items and total scores were significantly associated with violence ($p < .05$). With respect to the temporal proximity model, 9 of the 10 HCR-20 C and R items, and subscale scores were significantly associated with violence ($p < .05$). In addition, 11 of the 17 SAPROF items and total scores were significantly associated with violence ($p < .05$).

Third, the authors examined predictive accuracy of the HCR-20 Clinical and Risk Management subscales, and the SAPROF using multilevel modeling controlling for age, gender, and baseline diagnosis. With respect to the lagged model, 7 of the 10 HCR-20 C and R items, and subscale scores were significantly associated with violence, as well as 13 of the 17 SAPROF items and SAPROF total scores ($p < .05$). With respect to the temporal model, 9 of the 10 HCR-20 C and R items of subscale scores, and 14 of 17 SAPROF items and total scores were significantly associated with violence ($p < .05$).

Last, the authors re-ran the above temporal models, simultaneously entering all HCR-20 and SAPROF items. Among the 10 dynamic items on the HCR-20, three items demonstrated independent effects on violent behavior: violent ideation, instability and poor stress/coping ($p < .05$). Among the items on the SAPROF, one item demonstrated an independent effect on violent behavior: self-control ($p < .05$). The authors also reported the results of analyses when baseline-adjusted and fully adjusted coefficients were used to estimate the extent to which the association between a risk/protective factor and violent outcome was accounted for by the explanatory variables in the prior model. Implications for clinical practice are discussed.


Summary

This study investigated the reliability and predictive validity of the HCR-20 V3 in a sample of 387 patients discharged from 32 medium secure units in England and Wales. Using a prospective cohort follow-up design, patients discharged between 2010 and 2011 were observed over a 12 month period following discharge. A majority of the sample was male (89.2%, n = 345), Caucasian (59.7%, n = 231), had a diagnosis of schizophrenia (66.3%, n = 256), and had a history of serious violence (79.8%, n = 305). Average age of the sample at the time of the study was 37.6 years ($SD = 9.7$).

The HCR-20 V3 was completed baseline (discharge from forensic facilities) and follow-up violence data was collected at 6 and 12 months post-discharge using clinical records and information obtained through interview with a
social supervisor and/or care coordinator who knew the patient well. Violence was defined and measured using the MacArthur Community Violence Instrument, using sources of official police data, case file review, and interviews of participants and collaterals. Of the sample, 14% (n = 54) committed an act of violence at 6-months and 23% (n = 89) committed an act of violence at 12 months. Using information obtained from police records only, these rates were 1.6% (n = 6) and 2.8% (n = 10), respectively.

Interrater reliability of the HCR-20 V3, based on a subset of 20 cases, was reported for the Total, H, C, and R scales, respectively, as follows: .92, .91, .90, and .93. The means of the patients that were violent at 6 and 12 months were 26.59 (SD = 5.03) and 25.75 (SD = 5.69) respectively, and the means of the patients that were non-violent were 21.80 (SD = 3.30) and 21.51 (SD = 6.22) respectively. Mean differences in HCR-20 Total scores between violent and non-violent offenders at 6 (p < .001) and 12-month follow-ups (p < .001) was significant.

The authors reported AUCs for violence of the subscales and total score of the HCR-20 V3 at 6 and 12 months. For 6 months post-discharge AUC values were .73, .63, .75, and .67 for the total, H, C, and R scales, respectively. For 12 months post-discharge AUC values were .70, .63, .71 and .63 for the Total, H, C, and R scales, respectively. Additionally, the authors provided the point biserial correlations (rpb) with frequency of violence at 6 and 12 months. For 6 months correlations between Total (rpb = .23; p < .001), H (rpb = .14; p < .01), C (rpb = .22; p < .001), and R (rpb = .18; p < .001) scores and frequency of violence were significant. For 12 months correlations between Total (rpb = .23; p < .001), H (rpb = .14; p < .01), C (rpb = .24; p < .001), and R (rpb = .19; p < .001) were also significant. No other correlations were significant (p > .05).

Last, the authors used a median split on the HCR-20 V3 total and subscales to calculate sensitivity, specificity, positive predictive values (PPV), and negative predictive values (NPV). Median values were 23, 13, 5, and 4 on the HCR-20 total, H, C, and R scales, respectively. At 6 month follow-up, values were as follows: HCR-20 Total (sensitivity % = 81.5, specificity % = 60.7, PPV % = 25.1, NPV % = 95.3), HCR-20 H (sensitivity % = 70.4, specificity % = 53.2, PPV % = 19.6, NPV % = 91.7), HCR-20 C (sensitivity % = 77.8, specificity % = 66.7, PPV % = 27.5, NPV % = 94.9) and HCR-20 R (sensitivity % = 74.1, specificity % = 55.3, PPV % = 21.2, NPV % = 92.9). At 12 month follow-up, values were as follows: HCR-20 Total (sensitivity % = 73.6, specificity % = 63.0, PPV % = 36.6, NPV % = 89.2), HCR-20 H (sensitivity % = 66.7, specificity % = 54.7, PPV % = 29.9, NPV % = 85.0), HCR-20 C (sensitivity % = 65.5, specificity % = 68.0, PPV % = 37.3, NPV % = 87.2) and HCR-20 R (sensitivity % = 66.7, specificity % = 56.3, PPV % = 30.7, NPV % = 56.3).

The authors concluded that in addition to good inter-rater reliability, the HCR-20 V3 discriminated well between violent and non-violent participants, was strongly associated with frequency of violence, and was moderately predictive of post-discharge violence. The authors note, however, that it was difficult to get agreement on relevant factors. Study limitations and recommendations for future research are provided.

See Also


Project and Scholarly Work


Summary

The current study was conducted to examine the concurrent and predictive validity of the HCR-20 V3 and Hamilton Anxiety of Risk Management: Forensic Version (HARM-FV) in a forensic psychiatric sample. The HARM-FV is a 14-item risk assessment instrument developed for use in forensic inpatient settings to assess change in dynamic risk factors and facilitate risk management. The study sample consisted of 39 participants found unfit to stand trial (8%, n = 3) or NCRMD (92%, n = 36). The HARM-FV was prospectively rated by clinical team members, whereas the HCR-20 was prospectively rated file by 5 independent trained raters. Each participant was assessed using the HARM-FV and HCR-20 monthly over a six month period. Outcomes were recorded using the Aggressive Incidents Scale (AIS) and the Modified Overt Aggression Scale (MOAS), which assess the presence of inpatient intimidation, threats, and physical violence. Interrater reliability of the HCR-20 was reported for a subset of 11 cases. At 1- and 3-month follow-ups, ICC for V3 Summary Risk Ratings was excellent (ICCs = .80).

First, the authors presented correlations between the HARM-FV and HCR-20 V3. HCR-20 V3 summary risk ratings (SRRs) – including Future Violence/Case
Prioritization, Serious Physical Harm, and Imminent Violence— which were significantly associated with HARM-FV Immediate and Short Term Professional Support at the 1- and 3-month follow-ups ($r = .50\text{ to } .62, p < .01$). Kappa coefficients were also reported for HCR-20 and HARM-FV. There was significant rater agreement between HCR-20 Future Violence/Case Prioritization ratings and HARM-FV Immediate and Short Term Professional Support at 1 and 3 months ($\kappa = .32\text{ to } .46, p < .001$).

Second, the authors examined whether aggressive incidents recorded by the AIS were related to MOAS scores. Results indicated a significant correlation between the two measures ($r = .92, p < .01$). In addition, there was substantial agreement between the two tools when scores were dichotomized to indicate the presence or absence of aggressive outcomes over the follow-up ($\kappa = .79, p < .001$).

Last, the authors examined the predictive accuracy of summary risk ratings on the HARM-FV for: (a) any aggression (AIS Scales 1-9), (b) aggression including intimidation and any physical aggression, including property damage (AIS Scales 2-9), and (c) any physical aggression, including property damage (AIS Scales 4-9) using ROC-AUC and logit logistic regression analyses. Logit logistic regression analyses indicated that SRRs of Immediate and Short-Term Without Professional Support on the HARM-FV were predictive of AIS 1-9 and AIS 2-9 ($p < .05$), whereas SRRs of Immediate and Short-Term With Professional Support were predictive of AIS 4-9 ($p < .05$). No other associations were significant ($p > .05$). AUCs ranged from .60 to .68. The authors did not report findings on the predictive validity of the HCR-20 V3. The authors concluded that there is clinical support for the use of the HARM-FV within an inpatient forensic setting.

**See Also**


**Project and Scholarly Work**


**Summary**

The individual assessment and management of terrorism threat is of critical importance globally. In recent years, two structured professional judgment (SPJ) tools were developed to aid professionals in the assessment and management of individual terrorist threat: (a) the Multi-Level Guidelines (MLG), which was developed for the assessment and management of various types of group-based violence, including terrorist violence; and (b) the second version of the Violent Extremism Risk Assessment (VERA-2), which was developed specifically for the assessment and management of extremist violence. In this study the authors examined the application of the MLG, VERA-2, and HCR-20 V3 to five terrorism cases. These cases were selected to be diverse in terms of offender motivation and roles, time periods, and countries.

First, the authors examined whether the MLG and HCR-20 V3 could be applied reliably to terrorism cases. With respect to the MLG, results indicated substantial agreement in presence ($\kappa = .80$) and relevance ratings ($\kappa = .70$) on individual MLG items, substantial to perfect agreement on domain scores ($\kappa = .70\text{ to } .83$), and fair to moderate agreement in terms of summary risk ratings (SRRs; $\kappa = .21\text{ to } .47$). With respect to the HCR-20 V3, results indicated moderate agreement in presence ratings ($\kappa = .43$) and fair agreement in relevance ratings ($\kappa = .28$) of individual HCR-20 V3 items, slight to substantial ($\kappa = .11$ to .74) agreement on subscale scores, and less than chance to slight agreement in terms of SRRs ($\kappa = .05$ to .05). The limits of conducting reliability analyses on only five cases must be pointed out.

Second, the authors examined agreement between the MLG and HCR-20 V3 with respect to future risk for violence. Future Violence/Case Prioritization ratings had almost perfect agreement ($\kappa = 1.00$), whereas Imminent Risk ($\kappa = .50$) and Life Threatening Risk Ratings ($\kappa = .00$) had moderate and less than change agreement, respectively.

Last, the authors conducted a content analysis of the MLG, VERA-2, and HCR-20 V3. Some content overlap between the MLG and VERA-2, but not the HCR-20 V3 was observed. The authors concluded that the VERA-2 and MLG may be useful for assessing and managing terrorist violence risk, and that the HCR-20 V3 may have limited applicability in cases where other motivational factors are present (e.g., mental health). Results were also discussed with general considerations for assessing terrorism threat among individuals within forensic mental health systems.
Summary

In this paper, the authors presented user feedback and empirical validation data of the HCR-20 V3 at a forensic psychiatric hospital in the Netherlands. First, the authors presented the results of a qualitative study examining the user friendliness and perceived value of the HCR-20 V3 in clinical practice in sample of 174 assessors who attended an initial workshop for a Dutch version of V3 in 2013, and 22 new assessors who had used the V3 for approximately 2 years following training in 2013. In both the 2013 and 2015 samples, participants found that V3 was slightly harder to and took longer to code than V2. More than half of the 2013 and 2015 samples thought the following components of V3 were useful: sub-items (89% and 86%), indicators (78% and 77%), relevance (75% and 52%), scenarios (74% and 91%), and additional final judgments (67% and 27%). General impressions were that the V3 was an improvement over the V2 in terms of value, structure, comprehensiveness, and clinical utility. However, compared to the V2, V3 was more time intensive, sometimes unclear or complicated, required a lot of work, and more extensive. Overall, 99.4% (n = 173) and 100% (n = 22) of participants in the 2013 and 2015 samples thought training on V3 was worth the investment. In both samples, > 95% of participants though the V3 provided additional value than V2.

Second, the authors presented findings from a prospective validation study of the HCR-20 V3 and Structured Assessment of Protective Factors for Violence Risk (SAPROF) in a sample of 197 male and female inpatients at different stages of treatment. Internal consistency and interrater reliability of the HCR-20 V3 and SAPROF were not reported. A majority of the sample was male (75%, n = 148). Of the sample, 55% (n = 108) were intramural or on supervised leave (i.e., Inside Group) and 45% (n = 89) were on unsupervised leave or transmural (i.e., Outside Group). Violent incidents during treatment were recorded over a 9 to 12 month follow-up period. Of the sample, the overall incident of serious violence was 12% (n = 23). Compared to patients in the Outside Group, patients in the Inside Group had higher rates of serious violence (23% vs. 5%). The correlation between the HCR-20 V3 and SAPROF was $r = -.65$. Patients in the Inside Group had higher V3 and combined HCR-SAPROF scores, and lower SAPROF scores than patients in the Outside Group (means and standard deviations not reported).

The authors reported AUCs for the overall sample, and separately for each subgroup. In the total sample, AUCs were .68, .69, .69, and .69 for HCR-20 V3 total score, SAPROF total score, combined HCR-SAPROF total score, and combined HCR-SAPROF final risk judgment, respectively. In the Inside Group, AUCs were .52, .52, .52, and .65 for HCR-20 V3 total score, SAPROF total score, combined HCR-SAPROF total score, and combined HCR-SAPROF final risk judgment, respectively. In the Outside Group, AUCs were .80, .77, .80, and .64 for HCR-20 V3 total score, SAPROF total score, combined HCR-SAPROF total score, and combined HCR-SAPROF final risk judgment, respectively.

The authors also reported AUCs separately by gender. For males, AUCs were .71, .78, .76, and .68 for HCR-20 V3 total score, SAPROF total score, combined HCR-SAPROF total score, and combined HCR-SAPROF final risk judgment, respectively. For females, AUCs were .64, .56, .60, and .71 for HCR-20 V3 total score, SAPROF total score, combined HCR-SAPROF total score, and combined HCR-SAPROF final risk judgment, respectively.

The authors concluded that the HCR-20 V3 had been implemented successfully in clinical practice. The authors also noted that poorer predictive utility for patients in the Inside Group compared to patients in the Outside Group may be the result of improved risk management in the hospital rather than a limitation of the HCR-20 V3 or SAPROF.

Summary

In this paper, the authors presented implementation and clinical experiences using a preliminary version of HCR-20 V3 in forensic psychiatric practice in the Netherlands. First, the authors present results of beta-testing with Draft 1 of the HCR-20 V3. Seven researchers/diagnosticians and three treatment supervisors assessed one or more different forensic psychiatric patients with Draft 1. Raters reported several positive aspects (e.g., improved terminology, enhanced relevance to risk management). However, critical feedback was also reported (e.g., more complex coding scheme, ambiguity in some of the item descriptions). Overall, Draft 1 was viewed as a valuable revision of the HCR-20, but further improvements and clarifications were needed.

Second, the authors reported findings from a pilot study using a draft of the V3. Using a retrospective file study, 86 discharged patients were assessed using the HCR-20 and...
HCR-20 V3. Interrater reliability using intraclass correlation coefficients (ICCs) was calculated on a subsample of 25 cases and was good for both draft of V3 total score and the Future Violence/Case Prioritization summary risk rating (SRR) (ICC = .84 and .72, respectively). Correlation between draft V3 and HCR-20 was \( r = .93 \). Predictive validity of the HCR-20 and HCR-20 V3 were evaluated with respect to violent recidivism after discharge at one-, two-, and three-year follow-up using Receiver Operating Characteristic (ROC) analyses. Area under the curve (AUC) values of the HCR-20 V3 Total score (all \( ps < .05 \)) were as follows: .77 (one-year follow-up); .75 (two-year follow-up); and .67 (three-year follow-up). For the HCR-20 V2 Total score, AUC values were comparable and not significantly different (AUCs = .80; .74; and .67 at one-, two-, and three-year follow-ups, respectively). AUC values for the three-point V3 SRRs (all \( ps < .05 \)) were as follows: .72 (one-year follow-up); .67 (two-year follow-up); and .64 (three-year follow-up). Somewhat stronger results (\( p < .05 \)) were found for the five-point scale V3 SRRs (all \( ps < .05 \): AUCs were .82 (one-year follow-up); .74 (two-year follow-up); and .71 (three-year follow-up).

Third, the authors presented the results of a qualitative study examining the training friendliness and training necessity of the draft in a sample of 192 individuals who attended one of 11 workshops for a Dutch version of V3. Of those that attended, 68.8% (\( n = 132 \)) were sociotherapists and 25.5% (\( n = 48 \)) were psychologists or psychiatrists. More than half (55.2%, \( n = 106 \)) of the sample had more than one year experience with the HCR-20. On average, it took participants 50 minutes to code the HCR-20 and an average of 27 extra minutes to code V3. Participants found that there were easier sub-items on V3 compared to V2, however V3 items were harder to code because they were new. More than half of the sample thought the following components of V3 were useful: sub-items (88.6%, \( n = 163 \)), indicators (79.7%, \( n = 145 \)), relevance (76.1%, \( n = 140 \)), risk formulation (74.0%, \( n = 134 \)), scenarios (74.7%, \( n = 136 \)), and additional final judgments (67%, \( n = 118 \)). General impressions were that V3 was more applicable, clear, structured, detailed, more specific and more dynamic.

Last, the authors presented a case study and provide a description of the seven assessment and scoring steps of the HCR-20 V3 applied to the case. The authors concluded with the recommendation that the HCR-20 V3 be used for violence risk assessment and guidance for risk management. Directions for future research were also discussed.


**Summary**

In this paper, the authors present a clinical case study of a forensic psychiatric patient who was evaluated using the HCR-20 V3. First, the authors’ discuss reasons for the revision of the HCR-20, including the growing literature on violence risk and the need for clarification and improvement of the HCR-20 V2. Second, the authors present the facts of the case study and discuss the seven steps in completing the HCR-20 V3 (i.e., collecting case information, assessing the presence and relevance of risk factors, risk formulation, risk scenarios, management strategies, and conclusory opinions). The authors conclude that the HCR-20 V3 allows for more in-depth and specific assessment of an individual’s risk for future violence and leads to clear guidelines for risk management.

**See Also**


**Project and Scholarly Work**


**Summary**

The current study evaluated the interrater reliability and concurrent validity of the HCR-20 V3 among 32 forensic psychiatric patients in Sweden. Three sets of ratings on the HCR-20 Version 2 and Version 3 were completed by experienced clinicians on the basis of psychiatric files and patient interview for 32 patients.

Interrater reliability was assessed using intraclass correlation coefficients (ICCs; mixed, two-way; raters random, measure fixed, absolute agreement). With respect to presence ratings, results indicated excellent interrater reliability for HCR-20 total and subscale scores, and SRRs.
HCR-20 Review and Annotated Bibliography

Summary

This study compared V2 and V3 assessments of a male maximum-security patient. Two independent raters completed V2 and V3 together with hospital staff. Outcome of violence risk assessment, violence risk management plan, and clinical utility was discussed. The raters found that V2 provided less support in the assessment process and assisted with more broad and general risk management planning. In contrast, the V3 provided better structure for the assessment process. Further, although the V3 was more time consuming it was more comprehensive regarding the whole process of violence risk assessment.

Project and Scholarly Works


See Also


Project and Scholarly Work


ICC values were: H = .94, C = .86, R (In) = .69, R (Out) = .75, HCR-Total (In) = .94, HCR-20 Total (Out) = .94, Summary Risk Rating (In) = .81, and Summary Risk Rating (Out) = .75. ICC values were: H = .98, C = .95, R (In) = .87, R (Out) = .90, HCR-Total (In) = .98, HCR-20 Total (Out) = .98, Summary Risk Rating (In) = .93, and Summary Risk Rating (Out) = .90. With respect to relevance ratings, results indicated acceptable to excellent interrater reliability. ICC values were: H = .62, C = .60, R (In) = .74, R (Out) = .69, HCR-Total (In) = .76, and HCR-20 Total (Out) = .80. ICC values were: H = .83, C = .82, R (In) = .89, R (Out) = .87, HCR-Total (In) = .91, and HCR-20 Total (Out) = .93.

Interrater reliability of each of the individual HCR-20 items was also reported. With respect to presence ratings, ICC values ranged from .48 to .96 for H subscale items, .48 to .95 for C subscale items, and .15 to .88 for R subscale items. ICC values ranged from .72 to .99 for H subscale items, .71 to .98 for C subscale items, and .35 to 1.00 for R subscale items. With respect to relevance ratings, ICC values ranged from .00 to .83 for H subscale items, .29 to .78 for C subscale items, and .00 to .80 for R subscale items. ICC values ranged from .00 to .94 for H subscale items, .55 to .92 for C subscale items, and .00 to .92 for R subscale items. The authors determined that low ICC values (under .40 or below) were in large part due to little or no variance.

Concurrent validity between Version 2 and Version 3 was assessed via Pearson’s product moment correlation coefficients. The correlations were as follows (all rs significant at p < .001): H (.87), C (.76), R (In) (.67), and R (Out) (.82), Total Score (In) (.85), and Total Score (Out) (.90). Summary risk ratings on V2 were not available, thus concurrent validity between SPJ on V3 and V2 was not reported. Implications for the use of the HCR-20 V3 were discussed.

Concurrent validity between Version 2 and Version 3 was assessed via Pearson’s product moment correlation coefficients. The correlations were as follows (all rs significant at p < .001): H (.87), C (.76), R (In) (.67), and R (Out) (.82), Total Score (In) (.85), and Total Score (Out) (.90). Summary risk ratings on V2 were not available, thus concurrent validity between SPJ on V3 and V2 was not reported. Implications for the use of the HCR-20 V3 were discussed.

See Also


Project and Scholarly Work


The authors determined that low ICC values (under .40 or below) were in large part due to little or no variance.

Concurrent validity between Version 2 and Version 3 was assessed via Pearson’s product moment correlation coefficients. The correlations were as follows (all rs significant at p < .001): H (.87), C (.76), R (In) (.67), and R (Out) (.82), Total Score (In) (.85), and Total Score (Out) (.90). Summary risk ratings on V2 were not available, thus concurrent validity between SPJ on V3 and V2 was not reported. Implications for the use of the HCR-20 V3 were discussed.

Concurrent validity between Version 2 and Version 3 was assessed via Pearson’s product moment correlation coefficients. The correlations were as follows (all rs significant at p < .001): H (.87), C (.76), R (In) (.67), and R (Out) (.82), Total Score (In) (.85), and Total Score (Out) (.90). Summary risk ratings on V2 were not available, thus concurrent validity between SPJ on V3 and V2 was not reported. Implications for the use of the HCR-20 V3 were discussed.

See Also

aggressive incident (44.9%, \(n = 43\)) verbal aggression (17.3%, \(n = 16\)) physical aggression against property (2.0%, \(n = 2\)), physician aggression against others (38.8%, \(n = 37\)) and sexual aggression (5.1%, \(n = 5\)). Base rates of aggression at the median 34-month follow-up were as follows: any aggressive incident (33.7%, \(n = 33\)), verbal aggression (14.3%, \(n = 16\)), physician aggression against property (2.0%, \(n = 2\)), physician aggression against others (26.5%, \(n = 5\)) and sexual aggression (7.1%, \(n = 7\)). Mean scores, internal consistency, and interrater reliability of the HCR-20 V3 were not reported.

First, the authors examined the predictive utility of baseline HCR-20 V3 assessments for aggression over the median 34-month follow-up using Receiver-Operating Characteristics analyses. Area under the curve (AUC) values were .73 (95% CI = .63-.83, \(p < .001\)), .74 (95% CI = .63-.83, \(p < .001\)), and .61 (95% CI = .49-.72, \(p > .05\)) for the HCR-20 V3 H, C, and R scales, respectively.

Second, to examine whether dynamic risk factors changed over time, the authors calculated change scores for each pair of V3 assessments included in the study. Overall, decreased risk was observed for all individual Clinical and Risk Management items and subscale scores (\(p\) values not reported). However, greater risk was observed across all three assessments. In particular, the authors found that 68% (\(n = 66\)) of the scores changed on the HCR-20 R scale from evaluation 1 to evaluation 3, followed by 81% (\(n = 78\)) on the HCR-20 C scale and 85% (\(n = 82\)) on the HCR-20 Dynamic (i.e., combined C and R) scale.

Third, the authors examined the predictive utility of change scores between evaluation 1 and 2 and aggression at evaluation 3. The authors found that change in risk was not associated with change in aggression outcomes. AUCs were .55 (95% CI = .43-.67, \(p > .05\)), .49 (95% CI = .37-.61, \(p > .05\)), and .54 (95% CI = .42-.67, \(p > .05\)) for change in C, R, and Dynamic scales, respectively.

The authors noted that findings should be interpreted with caution given that participants had a wide range in length of hospitalizations prior to the start of the study that were not controlled for in the analyses. Further, participants were in a secure treatment facility, which could potentially decrease violence (that is, staff ostensibly could intervene if a patient showed exacerbation in risk factors).


The current archival study examined the incremental and predictive utility of the Female Addition Manual (FAM) and HCR-20 V3 among a sample of 28 female insanity acquittees admitted to a forensic psychiatric hospital in New York City. Six trained graduate research assistants completed the HCR-20 V3 and FAM using information provided in dangerous evaluations. Outcomes were collected from hospital incident reports and dangerousness evaluations and coded using the START Outcome Scale (SOS). Of the sample, 50% (\(n = 14\)) engaged in some act of aggression while in the hospital (i.e., 14.3% \([n = 5]\) verbal aggression, 10.7% \([n = 3]\) aggression against property, and 46.4% \([n = 13]\) physical aggression).

With respect to the HCR-20, mean scores were as follows: Total (27.32, \(SD = 6.53\)), Historical (15.11, \(SD = 3.05\)), Clinical (5.57, \(SD = 2.18\)), and Risk Management (6.64, \(SD = 2.93\)). With respect to the FAM, mean scores were as follows: Total (32.75, \(SD = 7.22\)), Historical (18.68, \(SD = 3.38\)), Clinical (6.43, \(SD = 2.70\)), and Risk Management (7.64, \(SD = 2.98\)). Inter-rater reliability (number of cases not reported) was also examined. ICC values ranged between .64 and .98 for HCR-20 Historical Items, .78 and .93 for HCR-20 Clinical Items, and .65 and .90 for HCR-20 Risk Management Items. ICC values could not be calculated for H1 (Violence), H3 (Relationships), H6 (Major Mental Illness) and C2 (Violent Ideation or Intent) due a lack of variance. ICC values ranged between .47 and .77 for FAM Historical Items, .47 and .72 for FAM Clinical Items, and .61 and .83 for FAM Risk Management Items. ICC for the combined HCR-FAM SRR was .91.

Correlation analysis was conducted to examine the association between risk assessment scores and aggression. With respect to any aggression, correlations were: HCR-20 V3 Total (.54, \(p < .01\)), FAM Total (.15, \(p > .05\)), HCR-20 Historical (.37, \(p > .05\)), FAM Historical (.49, \(p > .05\)), HCR-20 Clinical (.53, \(p < .01\)), FAM Clinical (.21, \(p > .05\)), HCR-20 Risk Management (.42, \(p < .05\)), and FAM Risk Management (.21, \(p > .05\)). With respect to physical aggression, correlations were: HCR-20 V3 Total (.49, \(p < .01\)), FAM Total (.16, \(p > .05\)), HCR-20 Historical (.35, \(p > .05\)), FAM Historical (.21, \(p > .05\)), HCR-20 Clinical (.49, \(p < .05\)), FAM Clinical (.13, \(p > .05\)), HCR-20 Risk Management (.36, \(p > .05\)), and FAM Risk Management (.29, \(p > .05\)).

ROC-AUC analyses were also conducted. With respect to any aggression, AUCs were: HCR-20 V3 Total (.81, \(p < .01\)), FAM Total (.83, \(p < .01\)), HCR-20 Historical (.72, \(p < .01\)), FAM Historical (.68, \(p > .05\)), HCR-20 Clinical (.83, \(p < .01\)), FAM Clinical (.82, \(p < .01\)), HCR-20 Risk Management (.75, \(p < .05\)), and FAM Risk Management (.80, \(p < .05\)). With respect to physical aggression, correlations were: HCR-20 V3 Total (.78, \(p < .05\)), FAM Total (.80, \(p < .01\)), HCR-20 Historical (.72, \(p > .05\)), FAM Historical (.68, \(p > .05\)), HCR-20 Clinical (.81, \(p < .01\)),

FAM Total (.80, \(p < .01\)), HCR-20 Historical (.72, \(p > .05\)), FAM Historical (.68, \(p > .05\)), HCR-20 Clinical (.81, \(p < .01\)),
Hierarchical logistic regression analyses were conducted to examine whether FAM total and subscale scores added incremental predictive utility to HCR-20 scores for the prediction of any aggression. No significant incremental effects were observed (p > .05). However, the authors noted that this may be due to the small sample size. Directions for future research were discussed.


Summary

The current retrospective study examined the interrater reliability and predictive validity of the HCR-20 V3 in a sample of 105 insanity acquittees in the USA. Mean age of the sample was 44.9 years (SD = 12.9). A majority of the sample was male (79%, n = 83) and African American (49%, n = 51). Of the sample, 77.1% (n = 81) were diagnosed with a psychotic disorder, 19.0% (n = 20) a mood disorder, 73.3% (n = 77) a substance abuse disorder, and 64.8% (n = 68) personality disorder or traits.

HCR-20 V3 assessments were completed on the basis of available file information by trained graduate student raters. Inpatient aggression was assessed over an average of 15.64 months (SD = 6.01, range: 4-24) with the START Outcome Scale (SOS) using information collected from hospital records. Of the sample, 56.2% (n = 59) had engaged in any aggression over the follow-up period. The most common type of aggression was physical aggression towards others (41.0%, n = 24), followed by verbal aggression (40.0%, n = 23), sexual aggression (12.4%, n = 7), and stalking (4.8%, n = 3).

First, the authors examined the interrater reliability (number of cases not reported) of HCR-20 V3 presence and relevance ratings using intraclass correlation coefficients (ICCs). With respect to presence ratings, ICCs were .82, .86, and .84 for HCR-20 Historical, Clinical, and Risk Management subscales, respectively. With respect to relevance ratings, ICCs were .82, .88, and .83 for HCR-20 Historical, Clinical, and Risk Management subscales, respectively. The authors also reported ICCs for individual presence and relevance ratings. With some exceptions, item ICC values were generally moderate to high (ICCs > .60).

Second, the authors examined the predictive validity of HCR-20 V3 assessments using Receiver Operating Characteristic (ROC) and logistic regression analyses. HCR-20 V3 Historical (AUC = .67, p < .01), Clinical (AUC = .78, p < .01), Risk Management (AUC = .64, p < .05), and Total scale scores (AUC = .75, p < .01) were significantly associated with any aggression over the follow-up. Controlling for length of follow-up and other subscale scores in logistic regression models, only HCR-20 V3 Clinical subscale scores remained significantly associated with any aggression over the follow-up (Exp [B] = 1.90, p < .01). The authors also reported the association between presence and relevance ratings on individual items and aggression over the follow-up.

Last, the authors examined the associations between presence and relevance ratings. Correlations between presence and relevance scales were as follows: HCR-20 V3 Historical subscale (r = .79, p < .01), Clinical subscale (r = .90, p < .01), Risk Management subscale (r = .86, p < .01), and Total scores (r = .88, p < .01). Limitations and directions for future research were discussed.

Project and Scholarly Works


Summary

The current study examined the concordance between HCR-20 V3 ratings of risk and patient perceptions of risk using data collected from 28 forensic psychiatric patients in Eastern Canada. The mean age of the sample was 39 years (SD = 10.74, range: 21-59). A majority of the sample was male (86%, n = 24), single (57%, n = 16), Caucasian (57%, n = 16), diagnosed with a psychotic (93%, n = 26) or comorbid substance abuse disorder (61%, n = 17), and had been convicted of a violent offence (82%, n = 23). Participants were asked to report two or three main factors that contributed to their index offense, as well as two or three main factors that would put them at risk for behaving violently in the future. For each HCR-20 V3 risk factors present in the participant’s response as contributing to either their index offense or future violence, a match was coded if a professional coded the same risk factor present.

First, the authors examined participant perceptions of risk. On average participants reported 2.2 risk factors (range: 0-5). Of the sample, 39% (n = 11) reported that they were not a future risk, 7% (n = 2) denied the index offense, and 7%
(n = 2) reported that they did not know what puts them at risk.

Second, the authors examined the degree to which patient perceptions of risk were concordant with those of HCR-20 V3 ratings. On average patients reported 1.5 risk factors (SD = 1, range: 0-3) that were consistent with those rated by a professional as present on the HCR-20 V3. Of the sample, 20% (n = 6) had 0 matches, 25% (n = 7) had 1 match, 40% (n = 11) had 2 matches, and 15% (n = 4) had 3 matches. The most commonly matched items were H6 (Major Mental Disorder; 46%) and C3 (Active Symptoms of Mental Disorder, 46%), followed by H10 (Treatment or Supervision, 25%), R5 (Stress or Poor Coping, 29%), C5 (Treatment or Supervision, 25%), R4 (Treatment or Supervision, 25%), H5 (Substance Abuse, 21%), and H3 (Relationships, 18%). Non-HCR risk factors reported by clients included self-defense, problematic social interactions, problematic thinking and choices, situational/circumstance factors, concrete interpretations, and delusional content (although some of these features could be considered HCR-related).

Third, the authors examined whether concordance between patient perceptions of risk and user risk ratings correlated with demographic, clinical, or legal variables. With respect to demographic variables, the authors found no associations between concordance and age, gender, education, work status, or previous psychiatric admissions (p > .05). However, for clients whom English was a primary language, ratings were more concordant compared to those with another primary language (p < .05). With respect to clinical variables, there was no association between concordance and psychiatric status, substance use, psychopathy, Violence Risk Appraisal Guide (VRAG) scores, or HCR-20 V3 Dynamic subscale scores (p > .05). However, there were significant correlations between concordance and HCR-20 V3 Total and Historical subscale scores, program compliance, and antisocial personality traits (p < .05). With respect to legal variables, there were no significant correlations between previous custody and perceived risk (p > .05). The authors concluded that engaging clients in the risk assessment is important because it can help clients to understand factors that increase risk.


Summary

The current study examined the predictive validity and clinical utility of modifiable risk factors for violence in a sample of 87 forensic psychiatric patients found Not Criminally Responsible on Account of Mental Disorder (NCRMD) transitioning to the community. Using a repeated-measures prospective design, theoretically based static and dynamic risk factors on the HCR-20 V3, Psychopathy Checklist-Revised (PCL-R), Brief Psychiatric Rating Scale (BPRS), State-Trait Anxiety Inventory (STAI), State-Trait Anger Expression Inventory (STAXI-2), Perceived Stress Scale (PSS), and Positive and Negative Affect Scale (PANAS) were assessed before hospital discharge, and in the case of dynamic measures, at 1 and 6 months post discharge. In addition, violence, victimization, and hospital readmission were measured at each community follow-up, and at 12 months post-discharge using police reports, health record information, and patient self-report (i.e., MacArthur Violence Risk Assessment Instrument, Self-Report of Delinquency).

Descriptive statistics of each of the measures were reported for the baseline assessment. At the baseline assessment, mean scores were as follows: BPRS Total (35.06, SD = 8.29), STAI Trait Anxiety (35.76, SD = 11.56), STAI State Anxiety (33.33, SD = 11.56), STAXI-2 Trait Anger (14.69, SD = 5.18), STAXI-2 State Anger (18.18, SD = 7.48), PSS Total (29.01, SD = 6.55), PANAS Positive Affect (34.21, SD = 8.15), PANAS Negative Affect (16.00, SD = 7.38), HCR-20 H (12.52, SD = 2.55), C (3.33, SD = 1.99), R (3.92, SD = 1.40), Total (19.78, SD = 4.40), and PCL-R total (15.26, SD = 6.42). Descriptive statistics of the dynamic measures was also reported at each of the follow-ups.

Interrater reliability of the HCR-20 V3 was calculated on a random sample of 12 cases. Interrater reliability was good to excellent for total and subscale scores (ICCs = .85 to .94 for baseline ratings and ICCs = .72 to .93 for follow-up 2), and good to excellent for summary risk ratings (ICC = 1.0 and .72 at baseline and follow-up 2, respectively). In addition, internal consistency was reported. At the baseline assessment, Cronbach’s alpha values were as follows: BPRS Total (.77), STAI Trait Anxiety (.91), STAI State Anxiety (.93), STAXI-2 Trait Anger (.90), STAI State Anger (.96), PSS Total (.75), PANAS Positive Affect (.87), PANAS Negative Affect (.92), HCR-20 H (.57), C (.68), R (.33), and Total (.71). Internal consistency was also reported at each of the follow-ups (α = .70 to .96 at follow-up 1 and α = .60 to .97 at follow-up 2).

First, dependent samples t-test was conducted to examine changes in dynamic measures over the follow-up. Results indicated significant declines in BPRS Total, PANAS Positive Affect, and PANAS Negative Affect between baseline and follow-up 2 (p < .05). No other differences were significant (p > .05). When the sample was median split according to static risk on the HCR-20 V3 H subscale, there were significant declines observed in the C and R subscales of the HCR-20V3, but only within the low static risk group (p < .05). No other differences emerged (p > .05). With respect to the summary risk rating of the HCR-20 V3, there was a significant increase in the proportion of
persons with serious mental illness (SMI) engage in

Second, cluster analyses was used to examine whether subgroups of individuals could be identified using dynamic measures that exhibited varying amounts and trajectories of change on the specified risk indicators, and whether these groups differed on the likelihood of adverse community outcomes. A two-cluster solution was obtained. Results indicated that participants in Cluster 1 had significantly lower levels of psychiatric symptoms, state anxiety and anger, negative affect and stress compared to Cluster 2 (p < .05). In addition, participants in Cluster 1 had significantly higher levels of positive affect at baseline (p < .05). Compared to Cluster 1 participants, Cluster 2 participants had higher rates of violence perpetration, violent victimization, and hospital readmission compared to Cluster 1 (p < .05). No other significant differences were found (p > .05).

Third, multinomial ordinal logistic regression was conducted to examine the incremental predictive efficacy of static and dynamic risk indicators measured at baseline for violence perpetration, victimization and re-hospitalization. First, whether dynamic risk factors added to the capacity of the Historical (H) subscale on HCR-20 V3 to predict outcomes was examined. Neither symptom- nor affect-based dynamic risk indicators accounted for additional variance in these models above and beyond H subscale scores (p > .05). In addition, neither C nor R subscale scores accounted for additional variance in these models above and beyond H subscale scores (p > .05). A similar pattern of findings was obtained when the PCL-R was used as an indicator of static risk.

Last, the ability of change in dynamic risk indicators to predict adverse community outcomes was investigated using Cox regression survival analyses with time-dependent covariates. Controlling for HCR-20 V3 H scale scores, change in HCR-20 V3 C and R scale scores were significantly associated with increased odds of engaging in violence (p < .05) and being re-hospitalized (p < .001). Of the symptom- and affect-based scales, only change scores on the BPRS had a significant association with victimization outcomes after controlling for the effect of static risk (p < .01).


Summary

Prior research has identified different motivations for why persons with serious mental illness (SMI) engage in offending. However, assessments of motivation have not typically been incorporated in structured instruments. To address this limitation, the Motivations for Offending in NCRMD Individuals at CAMH (MONICA) was developed. The MONICA examines the presence/absence of psychiatric symptoms and motivation influences prior to or during the index offense. The current study was conducted to validate the construct validity of the MONICA, and examine whether motivational classes on the MONICA were associated with life course trajectories of SMI and offending.

Study data were derived from 60 forensic psychiatric patients in Canada. A majority of the sample was male (90%, n = 54) and diagnosed with a psychotic disorder (91.7%, n = 55). The MONICA and HCR-20 V3 were coded on the basis of file review and a semi-structured interview. In addition, whether participants were early starters (i.e., age at first arrest or charge occurred prior to age of first psychiatric admission) or late starters (i.e. age at first arrest or charge occurred at the same time or after age of first psychiatric admission) was determined.

Mean scores on the HCR-20 V3 were as follows: Historical Presence (14.20, SD = 3.11), Clinical Presence (5.42, SD = 2.51), Risk Management Presence (4.72, SD = 1.76), Total Presence (24.33, SD = 5.51), Historical Relevance (12.07, SD = 3.30), Clinical Relevance (6.25, SD = 2.35), Risk Management Relevance (5.03, SD = 1.78), and Total Relevance (23.35, SD = 5.02). Prior to or during the index offence, 62.1% (n = 37) of the sample experienced hallucinations, 23.7% (n = 14) command auditory hallucinations, 100.0% (n = 60) delusions, 40% (n = 24) symptoms of a major mood disorder, and 75% (n = 45) substance use. Motivations for the index offence included: reactive anger (25%, n = 15), lust (11.7%, n = 7.0), financial gain (3.3%, n = 2), antisocial attitudes (11.7%, n = 7), poverty (3.3%, n = 2), and intoxication in the absence of psychosis (5.0%, n = 3). Of the sample, 53.3% (n = 32) were classified as early starters and 46.7% (n = 28) as late starters.

Patients with more delusional/hallucinatory motives for offending were characterized by a late starter life-course trajectory, less complex psychopathology, less substance use, a violent or sexual index offence, visual hallucinations prior to or during the index offense, great potentiality for more severe injury to the victim, and inability to understand the moral wrongness of their actions (p < .05). Patients with disorganized motives for offending were characterized by an early starter life course trajectory, more complex psychopathology, more substance use, less personality pathology, a non-violent index offense, less violent attitudes, more treatment/ supervision response problems, and judged to be at higher risk of future violence on the HCR-20 V3 (p < .05). Participants with more disorganized motives for offending tended to be characterized by more cluster B personality pathology, more comorbidity, non-
violent index offence, less severe victim injury, more relationship problems, more violent attitudes, more treatment/ supervision response problems, and higher HCR-20 V3 total scores ($p < .05$). The authors concluded that results support the clinical literature regarding subgroups of mentally disordered offenders, but have further implications regarding motivations for offending and risk in each of these groups.

### Project and Scholarly Work


### Summary

In this pseudo-prospective study, the authors examined the contribution of dynamic risk assessment measures to the prediction of inpatient violence in a sample of 99 forensic psychiatric patients. Average age of the sample at time of admission was 36.7 years ($SD = 12.8$). A majority of the sample was male (86%, $n = 85$), Caucasian (61%, $n = 60$), found NCRMD (52%, $n = 51$), diagnosed with a psychotic (61%, $n = 60$) or substance use disorder (70%, $n = 69$), and had committed a violent index offence (71%, $n = 70$). Three dynamic risk assessment measures were coded from archival data – the Violence Risk Scale (VRS), Short-Term Assessment of Risk and Treatability (START), and the HCR-20 V3. A static risk assessment measure was also coded (the Violence Risk Appraisal Guide-Revised or VRAG-R), as was a measure of psychopathy, the Psychopathy Checklist-Revised (PCL-R). Dynamic risk assessment tools were also re-coded using available information at discharge or using a cut-off date of December 31, 2013 if the patient was not discharged. Inpatient aggression was collected from file review using the Staff Observation Aggression Scale-Revised (SOAS-R). For the purpose of this study, violence was defined as any verbal, nonverbal, or physical behavior that was threatening to self, others or property, or physical behavior that actually caused harm (to self, others, or property). The base rate of any inpatient violence was 30% ($n = 29$).

Internal consistency was assessed using intra-class correlation coefficients (ICCs) on a randomly selected sample of 10 cases. ICC of the measures was as follows: HCR-20 V3 Total and Subscale scores (ICCs = .73-.86), PCL-R Total Score (ICC = .95), START Total and Subscale scores (ICC = .84-.89), and VRS (ICC = .94-.95). The VRAG was scored by only one coder and thus inter-rater reliability was not available.

First, the authors examined convergent validity between the HCR-20 and additional measures at the baseline assessment. HCR-20 Total Presence, Total Relevance, and SPJ Future Violence/Case Prioritization ratings were significantly associated with START Vulnerability, START Strength, VRS Total, and PCL-R total scores ($r = -.42$ to .81, $p < .05$). HCR-20 V3 Imminent Violence Ratings were significantly associated with START Vulnerability, START Strength, VRS Total, and VRAG Total ($r = -.25$ to .57, $p < .05$), but not PCL-R Total scores ($r = -.14$, $p > .05$). Correlations between HCR-20 Serious Physical Harm Ratings and START Vulnerability, START Strength, VRS Total, and PCL-R total scores were not significant ($r = -.14$ to .16, $p > .05$). Correlations between HCR-20 subscales and additional measures, as well as HCR-20 scales and subscales of the additional measures were also reported.

Second, the authors conducted repeated measures t-tests to examine change in risk assessment scores over time. The authors found significant changes in HCR-20 Clinical Presence and Relevance, HCR-20 Risk Management Presence and Relevance, START Vulnerability, START Strength, and VRS Dynamic scales ($p < .01$). No significant changes on HCR-20 Historical Presence and Relevant ratings ($p > .05$) or VRS Static scale scores were found ($p > .05$).

Third, the authors examined predictive utility of the risk assessment instruments using ROC-AUC analysis. AUCs for any inpatient violence were as follows: HCR-20 Total Presence (AUC = .76, $p < .001$), Historical Presence (AUC = .64, $p < .05$), Clinical Presence (AUC = .76, $p < .001$), Risk Management Presence (AUC = .76, $p < .001$), Total Relevance (AUC = .70, $p < .001$), Historical Relevance (.64, $p < .001$), C Relevance (.72, $p < .001$), R Relevance (AUC = .72, $p < .001$), Future Violence/Case Prioritization (AUC = .68, $p < .01$), Serious Physical Harm (AUC = .44, $p > .05$), and Imminent Violence (AUC = .75, $p < .001$), START Vulnerability Pre Treatment (AUC = .76, $p < .001$), Strength Pre-treatment (AUC = .69, $p < .01$), VRS Static Pre-treatment (AUC = .58, $p < .05$), VRS Dynamic Pre-treatment (AUC = .69, $p < .01$), VRAG (AUC = .60, $p > .05$), PCL-R Total (AUC = .63, $p < .05$), Facet 1 (AUC = .52, $p > .05$), Facet 2 (AUC = .62, $p < .05$), Facet 3 (AUC = .63, $p > .05$), Facet 4 (AUC = .66, $p < .05$), Factor 1 (AUC = .60, $p > .05$), and Factor 2 (AUC = .65, $p > .05$).

Fourth, the authors conducted hierarchical logistic regression models to examine the incremental validity of dynamic measures for inpatient aggression. In these models static (baseline) score for each measure and length of hospital stay were entered in the first step followed by the dynamic measure in the second step. Controlling for static scores and length of hospital stay, HCR-20 Clinical Presence (Exp[B] = 1.47, $p < .001$), HCR-20 Clinical Relevance (Exp[B] = 1.37, $p < .01$), and VRS Dynamic scale scores (Exp[B] = 1.11, $p < .01$) were significantly
associated with any inpatient aggression. No other incremental effects were observed ($p > .05$). The authors also report the findings of the analyses controlling for length of stay, VRAG, and static scores of the instrument.

Last, the authors conducted hierarchical logistic regression models to examine whether measured changes in risk measures would demonstrate incremental relationships with decreased inpatient aggression controlling for baseline total scores or scores on the respective scale. Several HCR-20 V3 change scores demonstrated incremental relationships with decreased inpatient aggression controlling for their respective baseline scores: Clinical Presence change scores, combined Clinical and Risk Management Presence change scores, Clinical Relevance change scores, and combined Clinical and Risk Management Relevance change scores ($\text{Exp} \ [B] < 1.00, p < .05$). In addition, several HCR-20 V3 change scores demonstrated incremental relationships with decreased inpatient aggression controlling for baseline HCR-20 V3 total scores: combined Clinical and Risk Management Relevance change scores and Risk Management Relevance change scores ($\text{Exp} \ [B] < 1.00, p < .05$). In addition, START Vulnerability ($\text{Exp} \ [B] = 0.88, p < .05$) and VRS Dynamic change scores ($\text{Exp} \ [B] = 0.35 \ p < 0.05$) each demonstrated incremental concurrent associations with decreased inpatient aggression. No other incremental effects were significant (all $ps > .05$).

The authors concluded that risk assessments using the HCR-20 V3, START, and VRS may inform the management and reduction of inpatient aggression, as well as assessments of dynamic risk more generally.

See Also


Project and Scholarly Work


Summary

Since the mid-1990s the HCR-20 has been implemented in the daily routines of the Haina Forensic Psychiatric Hospital in Germany. This paper provides an overview of the HCR-20 Version 3 workgroup established in the hospital, a description of the group’s work, and a summary of the main results of the workgroup.

This workgroup was established to systematically work on certain goals (e.g., giving user feedback on the draft version). Projects completed by the workgroup included beta-testing of the English HCR-20 V3 draft version, the German translation of chapters 4 (“Using the HCR-20 Version 3”) and 5 (“HCR-20 Version 3 Risk Factors and Coding Instructions”) of the HCR-20 User Manual, and an empirical evaluation of the German HCR-20 V3 draft version.

First, the authors describe beta-testing of the HCR-20 Version 3. Fifteen raters (experienced and inexperienced with using SPJ tools to evaluate risk for violence) rated one case each with the English draft version of the HCR-V3, following which they all completed a user feedback form. The raters’ cases had range of diagnoses (e.g., psychosis, personality disorder, and intellectual disability), level of familiarity (patient known vs. unknown to the rater) and security levels (no levels vs. unescorted levels). The user feedback form included the possibility to give item-by-item and overall feedback. Subsequently the data were qualitatively analyzed by the workgroup. Feedback from these raters was incorporated in revisions to V3.

Next, the authors describe an empirical evaluation of the interrater reliability of the HCR-20 Version 3 draft version. Fiver raters (post graduate students with a degree in psychology) each rated the presence and relevance of the 30 case vignettes of patients at the hospital with the German draft version of the HCR-20 V3. For the R ratings the raters were asked to only do community out ratings (i.e. as if the institutionalized person were to be released at the present moment with the existing plans). Of the patients illustrated in the case vignettes, 10 had psychosis, 10 had personality disorder, and 10 had an intellectual disability. Raters did not discuss cases and ratings with one another. The interrater reliability of the Future Violence/Case Prioritization summary risk rating was excellent, at ICC = .86. The interrater reliability of the 25 individual items and sub-items of the Historical subscale were between .06 (H3: Relationships) and .99 (H5: Substance Use; H6a: Psychotic Disorders). It was not possible to calculate the ICC for H1 (Previous Violence) due to the lack of variance between the different cases, since all evaluated patients had a violent index offence. ICC values for the 16 items and sub-items of the Clinical subscale were between .43 (C2: Violent Ideation or Intent) and .92 (C3a: Symptoms or Psychotic Disorders) ICC values for the 7 items and sub-items of the Risk Management subscale ranged between .51 (R3: Personal Support) and .86 (R4b: Treatment or Supervision Response – Responsiveness). ICC values of the summary risk ratings (of low, moderate, or high) were .86. The authors conclude with a discussion of possible reasons for
lower reliability (e.g., lack of variance, or misunderstanding from the training procedure) of some of the items.

See Also


Project and Scholarly Work


Summary

This research project examined the performance of a draft version of the Historical scale of the HCR-20 V3 based on archival data from 80 forensic psychiatric patients. Participants were randomly selected from a list of all insanity acquittees released (conditional discharge or visit leave) between 2000 and 2003 and returned to the forensic hospital. The average age of the patients at time of release was 35. Most participants were male (91.25%), Caucasian (76.3%), and had a previous psychiatric hospitalization (86.3%). For each participant, violence risk was coded on version 2 and version 3 of the Historical scale of the HCR-20, as well as on the VRAG.

The total scores on the revised HCR-20 Historical subscale displayed an increased range (7 to 30) compared to that of the total scores on the current version of the subscale (6 to 20). In addition, the former scores approached normality in their distribution, whereas the latter ones tended to be highly negatively skewed (56.4 % of the participants exhibited total scores between 18 and 20 on the existing version of the subscale).

Interrater reliability for the total scores on the three rated scales was examined over 12 pairs of ratings (3 raters) randomly chosen from the dataset. Interrater reliability of the sum of numerical presence ratings for the V3 historical factors was acceptable, ICC1 = .75, ICC2 = .85; and slightly higher than those mad using V2 historical risk factors ICC1 = .69 and ICC2 = .82. ICC values for the VRAG were ICC1= .79, ICC2= .88. The authors note that coding solely from files was challenging and may have affected interrater reliability. The authors recommend using both file and interview data, especially as it pertains to ratings made on version 3 of the HCR-20.

Version 3 of the H scale correlated significantly with both the VRAG (r = .60; p < .01) and Version 2 of the H scale (r = .60; p < .01). The correlation between the VRAG and Version 2 was slightly lower (r = .50; p < .01). Since Version 3 of the HCR-20 subscale correlated significantly with validated measures of historical risk factors for violence, the authors concluded that it evidenced concurrent validity.

Project and Scholarly Work


Summary

The study focused on changes in repeated assessments of the HCR-20 V2 and V3. The use of significant changes in clinical factors to inform discharge decisions was illustrated by a female forensic psychiatric patient case study. The authors found that sub-items of the V3 helped structure judgment. Further, presence and relevance ratings of helped define formulations and the specification of time frame and priority of case ratings helped enhance the structure of the risk management produce.
End of Forensic Psychiatric Settings
Project and Scholarly Work


Summary

There are a number of issues concerning the use of the HCR 20 V3 in pre-trial settings that have implications for the accuracy and utility of risk ratings and recommendations in this context. Common issues include limited file information, unwillingness or inability of the offender to provide accurate information during clinical interview, uncertainty among raters on how best to score items when the charge against the offender is not yet a conviction, and how to rate risk management items when the offender’s future living situation depends on trial outcome. The present study was conducted to investigate the use of the HCR-20 V3 among a sample of 84 male offenders in a pre-trial correctional facility in the USA. The HCR-V3 was coded on the basis of available interview and collateral data by four trained graduate research assistants.

Means on the HCR-20 V3 were as follows: Historical Presence (13.92, SD = 4.18), Clinical Presence (5.43, SD = 2.23), Risk Management (In) Presence (5.64, SD = 1.89), Risk Management (Out) Presence (6.98, SD = 2.32), Historical Relevance (10.88, SD = 5.28), Clinical Relevance (4.93, SD = 2.67), Risk Management (In) Relevance (3.63, SD = 2.73), and Risk Management (Out) Relevance (5.73, SD = 3.09). Descriptive statistics for each of the individual HCR-20 items were also reported.

To evaluate interrater reliability in the current study, 18% (n = 15) of cases were randomly selected to be independently rated by two interviewers. Interrater reliability was assessed using intraclass correlation coefficients (ICCs). ICCs for each sub-scale ranged from .92 (Historical) to .67 (Clinical). For the Risk Management subscale, ICCs were assessed for both incarcerated (.68) and community (.88) settings. ICC for Relevance ratings were higher for the Historical (.85) and Clinical (.77) subscales compared to the Risk Management scale, both for institutional (.48) and community (.67) settings.

Descriptive results from the present study were contrasted with those obtained from a Canadian inpatient psychiatric sample (n = 50) and a Canadian correctional sample (n = 56). Some group differences in risk ratings were observed (p values not reported). Ratings for the Historical scale among the present study sample were moderately higher than for the Canadian civil psychiatric sample (Cohen’s d = -2.00 to 1.56). Participants in the current study also evinced slightly greater Clinical scale scores than those obtained within the Canadian prison (Cohen’s d = -0.42 to 0.73) and psychiatric samples (Cohen’s d = -2.21 to 0.91). R-Community scale scores in the present study were greater than both the psychiatric (Cohen’s d = 0.28 to 2.26) and prison samples (Cohen’s d = 0.86 to 2.26).

The extent to which item Relevance ratings might incrementally contribute to predictive validity of the Presence Ratings in the prediction of summary risk ratings was examined using stepwise discriminant function analysis with the interaction between Relevance and Presence ratings as predictors. Results indicated that interaction scores uniquely predicted SRRs scales (p < .001). Implications for clinical practice, including potential modifications to the HCR-20 V3 that may be helpful or necessary in pre-trial settings, were discussed.

End of Correctional Settings
The present retrospective study investigated the concurrent validity of the HCR-20 Version 3 in a sample of 27 civil psychiatric patients and 16 offenders recruited from correctional institutions and probation offices in Western Canada. Violence was assessed retrospectively during a semi-structured interview and from file information. Information from these sources was used to determine an overall dichotomous perpetration of violence outcome and a dichotomous perpetration of physical violence outcome. The HCR: 20 V3 was coded using file information and interview. In this study, three decisions regarding the relevance of each risk factor were made: a general decision of relevance (Idiographic relevance), the relevance of this risk factor to past violence (Historical relevance), and the relevance of this risk factor to future violence perpetration (Future relevance).

ROC analyses were used to determine the ability of the HCR: V3 ratings systems to postdict the perpetration of any violence and physical violence. Overall, the AUCs were all larger for physical violence (average AUC = .75) compared to any violence (average AUC = .69). Comparing the different rating systems, the final risk judgments tended to yield the largest effects with AUCs of .76 for any violence and .83 for physical violence (average AUC = .79).

The study authors also performed postdictive analyses of the association between the Presence and Relevance ratings and violence. The presence scores yielded effects comparable to research on the former version of this instrument (HCR-20) with an average AUC of .67. Nearly all of the new rating schemes outperformed the presence scores. In decreasing order of effects, the historical relevance ratings yielded the largest effects (average AUC = .82), followed by the overall idiographic relevance ratings (average AUC = .75), and the indicator (manifestation) system (average AUC = .71). With regards to physical violence only, the future relevance ratings outperformed the presence scores; however, on average this rating scheme was slightly below the presence scores (average AUC = .66). Thus, the different rating schemes outperformed merely rating the presence of risk factors.

The authors conducted hierarchical logistic regression to assess the incremental validity of the different features in comparison to the presence ratings, the subscale presence scores were entered in the first block leading to a significant model. When the other scoring features were entered in the second block, individually, each of these features resulted in a significant increase in model fit. Thus, all the different features demonstrated incremental validity over the presence scores.

The present prospective study examined the predictive utility of the HCR-20 V3 in a mixed sample of 56 provincial offenders and 50 civil psychiatric patients who were transitioning out of institutions and into the community in Western Canada. Participants were included in the study if they were 18 years or older, fluent in English, and had no diagnosis of mental retardation. Most participants were male (59.4%, n = 63), Caucasian (74.5%, n = 79), and unemployed at time of entry into correctional or mental health systems (72.6%, n = 77). On average participants were 33.6 years of age (SD = 8.8). Many of the participants were diagnosed with a psychotic disorder or had significant psychotic features noted (33%, n = 35).

The HCR-20 Version 2 and 3 were coded by trained graduate student raters based on available file and interview data before participant release into the community or once recruited in the community, and up to five times post-discharge. Instances of violent behavior post-discharge were recorded after baseline interviews (i.e., over follow-up contacts 1 to 5) for a follow-up period varying from 1 (approximately 4–6 weeks on average) to 5 or more months (approximately 6–8 months on average). Consistent with the HCR-20 V3 manual violence was defined as any actual, attempted, or threatened infliction of bodily harm of another person (Douglas et al., 2013). The base rate of violence was
First, the authors examined mean scores on the HCR-20 for the combined sample and separately for psychiatric and offender subsamples. With respect to the combined sample, mean scores were as follows: HCR-20 H Presence (12.8, SD = 4.3), C Presence (4.9, SD = 2.5), R Presence (5.2, SD = 2.6), Total Presence (22.9, SD = 7.8), H Relevance (7.8, SD = 5.9), C Relevance (3.5, SD = 3.1), R Relevance (3.4, SD = 3.0), and Total Relevance (14.7, SD = 11.2). With respect to the psychiatric sample, mean scores were as follows: HCR-20 H Presence (10.9, SD = 4.3), C Presence (5.4, SD = 2.0), R Presence (4.0, SD = 2.3), Total Presence (20.3, SD = 7.2), H Relevance (6.3, SD = 5.7), C Relevance (3.8, SD = 3.3), R Relevance (2.5, SD = 2.6), and Total Relevance (12.6, SD = 11.0). With respect to the offender sample, mean scores were as follows: HCR-20 H Presence (14.5, SD = 3.6), C Presence (4.4, SD = 2.8), R Presence (6.3, SD = 2.5), Total Presence (25.2, SD = 7.7), H Relevance (9.2, SD = 5.8), C Relevance (3.3, SD = 2.8), R Relevance (4.2, SD = 3.2), and Total Relevance (16.6, SD = 11.1). HCR-20 total, H, and R scale scores were significantly higher among offenders compared to psychiatric patients (p < .05). No other differences were significant (p > .05).

Second, the authors tested whether gender or sub-sample type (i.e., offenders vs. civil psychiatric patients) impacted the predictive accuracy of the HCR-20 using multivariate logistic regression analyses. For both 4–6 week and 6–8 month follow-ups, gender did not moderate the association between HCR-20 V3 scores and violence (p > .05). With the exception of R scale relevance ratings (p < .05), sub-sample type did not moderate the association between. Thus, participants were collapsed for multivariate analyses.

Third, the authors examined concurrent validity between Version 2 and Version 3. Concurrent validity between the HCR-20 Version 2 and 3 was as follows (all ps < .001): H (r = .89), C (r = .76); R (r = .81), HCR-20 Total (r = .91), and SPJ (r = .98).

Fourth, the authors tested the predictive utility of the HCR-20 V3 using point biserial correlation coefficients and Receiver Operating Characteristic (ROC) analyses. All summary risk ratings (SRRs) were significantly associated with violence (all ps < .05). Correlations for SRRs ranged from .33 to .53, and AUCs ranged from .68 to .91. In the combined sample analyses, all Presence rating sums (i.e., H, C, R, and HCR-20 Total) were significantly associated with violence (all ps < .05). Correlations ranged from .26 to .46, and averaged .36. AUC values ranged from .71 to .78, and averaged .74. Similarly, for the combined sample analyses, all Relevance rating sums were significantly associated with violence (all ps < .05). Correlations ranged from .21 to .32, and averaged .27. Area under the curve (AUC) values ranged from .63 to .72, and averaged .68. Point biserial correlations and AUC values were also reported separately for each subsample and follow-up length.

Fifth, the authors conducted a series of logistic regression analyses to test (a) whether Relevance ratings added incrementally to Presence Ratings, (b) whether SRRs added incrementally to Presence Ratings, and (c) whether SRRs added incrementally to Relevance Ratings. The outcome variables in these analyses were violent offending during the 4–6 week follow-up and violent offending during the 6–8 month follow-up, respectively. Results indicated that SRRs added incrementally to both Presence and Relevance Ratings for the prediction of violent behavior over the two follow-up periods (p < .05). However, Relevance Ratings did not add incrementally to presence ratings for the H, C, and R and Total Scales (p > .05).

Last, the authors examined whether HCR-20 V3 indices were moderated by gender. No significant moderation effects were observed (p > .05). Finally, the authors examined correlations between the two versions of the HCR-20 and violence. Associations with violence were comparable between instruments (all ps < .001). The authors concluded that HCR-20 V3 had good validity among psychiatric patients and offenders released into the community. Directions for future research are discussed.

See Also


End of Mixed Settings
Scholarly Works that were Summarized in Section 1
(followed by the page number on which they appear)


V3. International Journal of Forensic Mental Health, 13(2), 140-147. ...................................................22


---

**Other Scholarly Works Cited in Section 1**


Awaiting Copies of the Following Presentations


Waerp J., Bell, C., & Tanke Gerd, H. (2016, June). Clinical experiences with Version 3 of the HCR-20, both from a forensic unit and from a general psychiatric unit. Poster presented at the annual meeting of the International Association of Forensic Mental Health Services, New York, USA.
HCR-20 V3 Related Disseminations and Academic Commentary

HCR-20 V3 Professional Manuals


HCR-20 V3 Risk Factor Literature Review and Rationale


HCR-20 V3 Translations

Translated into Dutch as:


Translated into Swedish as:


Related HCR-20 V3 Chapters and Articles by HCR-20 Authors


Published HCR-20 V3 Peer Academic Commentary


de Apodaca, R. F., Blake, B. T., & Stevens, L. P. (2015). Risk assessments in violence and sex cases: Current, practical advice for attorneys employing mental health


Section 2 | HCR-20\textsuperscript{V1/V2} Studies
Description of the HCR-20 V1/V2 Violence Risk Assessment Scheme

Last Updated January 1, 2014:

The Historical-Clinical-Risk Management-20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997) belongs to the Structured Professional Judgment (SPJ) model of violence risk assessment, along with instruments such as the SARA, RSVP, SVR-20, SAVRY, and START. It is intended to be used to guide a comprehensive, structured assessment of violence risk for adults (18+) within forensic psychiatric, civil psychiatric and offender samples. The conceptual scheme of the HCR-20 aligns risk factors into past, present, and future. Its 10 Historical factors obviously concern the past. However, the HCR-20 contains 5 Clinical items that are meant to reflect current, dynamic (changeable) correlates of violence. The future is recognized in the 5 Risk Management items, which focus attention on situational post-assessment factors that may aggravate or mitigate risk. These are also dynamic. The HCR-20 takes its name from these three scales — Historical, Clinical, Risk Management — and from the number of items (20). Table 2 shows the items.

Ultimately, the HCR-20, as with all SPJ instruments, is intended to establish the presence and individual relevance of important violence risk factors, and to inform the selection and intensity of risk management strategies. The HCR-20 is not an actuarial instrument, in that decisions about risk are not based on algorithms, equations, cut-offs, or other mechanical strategies. Rather, clinicians make structured professional judgments about risk level and degree of necessary management (low, moderate, or high). The procedure used by the HCR-20, and other SPJ instruments, is well-studied, research-based, and empirically-validated. It is also intended to be as clinically useful and informative as possible. As such, the HCR-20 is an attempt to merge science and practice by offering an instrument that can be integrated into clinical practice but also is empirically based and testable.

The HCR-20 was developed from a thorough consideration of the empirical literature concerning factors that relate to violence. It attempts to develop professional standards regarding the process and substance of risk assessments. Further, the HCR-20 integrates the experience of clinicians, and is easy to administer, understand, and score. Randy Borum (1996) has written about the HCR-20 that “the promise of this instrument lies in its foundation on a conceptual model or scheme for assessing dangerousness and risk; its basis in the empirical literature; its operationally defined coding system...[and] its practical use....The field eagerly awaits new data on this instrument” (p. 950). The field has changed since Borum wrote those words – there are now 233 disseminations on the HCR-20 reviewed in this document.
Table 4. Items in the HCR-20V1/V2 Violence Risk Assessment Scheme

<table>
<thead>
<tr>
<th>Sub-Scales</th>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Historical Scale</strong></td>
<td></td>
</tr>
<tr>
<td>H1</td>
<td>Previous Violence</td>
</tr>
<tr>
<td>H2</td>
<td>Young Age at First Violent Incident</td>
</tr>
<tr>
<td>H3</td>
<td>Relationship Instability</td>
</tr>
<tr>
<td>H4</td>
<td>Employment Problems</td>
</tr>
<tr>
<td>H5</td>
<td>Substance Use Problems</td>
</tr>
<tr>
<td>H6</td>
<td>Major Mental Illness</td>
</tr>
<tr>
<td>H7</td>
<td>Psychopathy</td>
</tr>
<tr>
<td>H8</td>
<td>Early Maladjustment</td>
</tr>
<tr>
<td>H9</td>
<td>Personality Disorder</td>
</tr>
<tr>
<td>H10</td>
<td>Prior Supervision Failure</td>
</tr>
<tr>
<td><strong>Clinical Scale</strong></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>Lack of Insight</td>
</tr>
<tr>
<td>C2</td>
<td>Negative Attitudes</td>
</tr>
<tr>
<td>C3</td>
<td>Active Symptoms of Major Mental Illness</td>
</tr>
<tr>
<td>C4</td>
<td>Impulsivity</td>
</tr>
<tr>
<td>C5</td>
<td>Unresponsive to Treatment</td>
</tr>
<tr>
<td><strong>Risk Management Scale</strong></td>
<td></td>
</tr>
<tr>
<td>R1</td>
<td>Plans Lack Feasibility</td>
</tr>
<tr>
<td>R2</td>
<td>Exposure to Destabilizers</td>
</tr>
<tr>
<td>R3</td>
<td>Lack of Personal Support</td>
</tr>
<tr>
<td>R4</td>
<td>Noncompliance with Remediation Attempts</td>
</tr>
<tr>
<td>R5</td>
<td>Stress</td>
</tr>
</tbody>
</table>

*Note. Adapted from Webster, Douglas, Eaves, and Hart (1997a).*
Receiver Operating Characteristic (ROC) Analyses

Last Updated September 22, 2010

Information on ROC analyses was included in the first version of this document, and referred to as “an emerging technique” in risk assessment. It is no longer “emerging,” but rather one of the most common means by which to evaluate the predictive validity of risk assessment instruments. As such, we’ve retained this section for any readers who might not yet be familiar with ROC analyses.

ROC statistical analysis is summarized here because most HCR-20 (and other risk assessment instruments) studies use this analysis, and results are reported in terms of the statistical indexes that ROC produces. Although ROCs have been used in the area of radiology (Lusted, 1978), radar signal detection, and sensory psychology since the 1950s and 1960s (Metz, 1984), they were introduced into the area of violence risk assessment in the 1990s (Douglas, Ogloff, Nicholls, & Grant, 1999; Mossman, 1994a, 1994b; Rice & Harris, 1995; Rice, 1997). They are recommended in this area because they are less dependent on the base rate of the criterion variable in the sample (in the present case, violence) than are traditional measures of predictive accuracy derived from 2 x 2 contingency tables (such as false positives and false negatives). Since correlations diminish with departures from base rates of 50%, correlational techniques are not the most effective means to estimate predictive efficiency of risk assessment schemes (Rice & Harris, 1995).

ROCs allow for the comparison of various thresholds on the predictor measures for offering predictions of violence, an overall index of accuracy which accounts for all possible thresholds, the simple identification of the optimal threshold, and the comparison of two or more predictors (Hsiao, Bartko, & Potter, 1989; Lusted, 1978; Metz, 1984; Mossman, 1994a; 1994b; Mossman & Somoza, 1991; Vida, 1997).

The term “receiver operating characteristic” took its name because it describes the detection, or prediction, “characteristics” of the test, and the “receiver” of the data can “operate” at any given point on the curve (Metz, 1978). ROCs are meant to be applied to data that are comprised of a continuous predictor variable and a dichotomous dependent measure. They take the form of a figure (see Sample ROC, next page, for an example) with the sensitivity (true positive rate [TPR]) of the predictor plotted as a function of the false positive rate (FPR [1-specificity]) (Mossman & Somoza, 1991). For any given level of specificity, the receiver knows the sensitivity. Each point on the curve (which corresponds to a cut-off on the predictor) represents a different trade-off between sensitivity and specificity.

The area under the curve (AUC) of the ROC graph can be taken as an index for interpreting the overall accuracy of the predictor. Areas can range from 0 (perfect negative prediction), to .50 (chance prediction), to 1.0 (perfect positive prediction). A given area represents the probability that a randomly chosen person who scores positive on the dependent measure (in this study, is actually violent) will fall above any given cut-off on the predictor measure, and that an actually non-violent person will score below the cut-off (Mossman & Somoza, 1991). Thus, an area of .75 means that there is a 75% chance that an actually violent person would score above the cut-off for violence on the predictor, and an actually non-violent person would score below the cut-off. AUC values of 0.70 may be considered moderate to large, and .75 and above may be considered large.
Figure 1. A Sample ROC Curve

![A Sample ROC Curve Diagram](image)

- **True Positive Rate (Sensitivity)**
- **False Positive Rate (1 – Specificity)**
- **Possible Operating Points**
- **The Line of No Information (Chance Prediction)**
- **All grey area = Area Under the Curve (AUC)**
Table 5. Summary of HCR-20\textsuperscript{V1/V2} Studies by Sample, Outcome, Gender, and Country

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Civil Psychiatric</th>
<th>Forensic Psychiatric</th>
<th>Correctional</th>
<th>Mixed</th>
<th>Juvenile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Entries</td>
<td>Total N</td>
<td>Entries</td>
<td>Total N</td>
<td>Entries</td>
<td>Total N</td>
</tr>
<tr>
<td>Total</td>
<td>237</td>
<td>42,013</td>
<td>28</td>
<td>4,045</td>
<td>125</td>
<td>17,841</td>
</tr>
<tr>
<td>By Outcome\textsuperscript{1}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Institution</td>
<td>76</td>
<td>11,065</td>
<td>12</td>
<td>1,763</td>
<td>52</td>
<td>7,334</td>
</tr>
<tr>
<td>Community</td>
<td>80</td>
<td>19,136</td>
<td>10</td>
<td>2,500</td>
<td>31</td>
<td>7,585</td>
</tr>
<tr>
<td>Both</td>
<td>31</td>
<td>3,399</td>
<td>4</td>
<td>410</td>
<td>17</td>
<td>1,812</td>
</tr>
<tr>
<td>By Sample Gender\textsuperscript{2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>136</td>
<td>27,523</td>
<td>23</td>
<td>3,672</td>
<td>76</td>
<td>11,955</td>
</tr>
<tr>
<td>Male</td>
<td>85</td>
<td>14,055</td>
<td>4</td>
<td>373</td>
<td>47</td>
<td>5,976</td>
</tr>
<tr>
<td>Female</td>
<td>6</td>
<td>435</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>90</td>
</tr>
</tbody>
</table>

\textsuperscript{1} The number of entries and total \( N \) do not always add up to match the totals row due to missing information from some studies. Also, some research projects were included in multiple rows due to the overall project collecting data regarding both inpatient and community recidivism, but particular studies that emerged from this project reporting on one time of outcome.

\textsuperscript{2} The number of entries and total \( N \) do not always add up to match the totals row due to missing information from some studies. Also, it is acknowledged that many of the studies that utilized mixed gender samples contained predominantly males.
<table>
<thead>
<tr>
<th>By Country</th>
<th>Civil Psychiatric</th>
<th>Forensic Psychiatric</th>
<th>Correctional</th>
<th>Mixed</th>
<th>Juvenile</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>60 13,914</td>
<td>6 1,079</td>
<td>36 6,737</td>
<td>12 4,999</td>
<td>6 1,099</td>
</tr>
<tr>
<td>Canada</td>
<td>49 6,371</td>
<td>8 410</td>
<td>24 3,643</td>
<td>7 855</td>
<td>8 1,251</td>
</tr>
<tr>
<td>Germany</td>
<td>22 2,675</td>
<td>-</td>
<td>8 965</td>
<td>6 1,241</td>
<td>5 469</td>
</tr>
<tr>
<td>USA</td>
<td>21 8,807</td>
<td>5 1,204</td>
<td>9 1,271</td>
<td>7 6,201</td>
<td>-</td>
</tr>
<tr>
<td>Sweden</td>
<td>21 1,913</td>
<td>-</td>
<td>16 1,489</td>
<td>3 186</td>
<td>2 238</td>
</tr>
<tr>
<td>Netherlands</td>
<td>17 1,714</td>
<td>-</td>
<td>17 1,714</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Belgium</td>
<td>7 922</td>
<td>-</td>
<td>4 619</td>
<td>1 68</td>
<td>2 235</td>
</tr>
<tr>
<td>Ireland</td>
<td>6 358</td>
<td>-</td>
<td>6 358</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Australia</td>
<td>5 501</td>
<td>-</td>
<td>3 208</td>
<td>1 59</td>
<td>1 234</td>
</tr>
<tr>
<td>International</td>
<td>4 533</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4 533</td>
</tr>
<tr>
<td>Portugal</td>
<td>3 316</td>
<td>-</td>
<td>-</td>
<td>3 316</td>
<td>-</td>
</tr>
<tr>
<td>Denmark</td>
<td>3 295</td>
<td>-</td>
<td>1 107</td>
<td>2 118</td>
<td>-</td>
</tr>
<tr>
<td>Spain</td>
<td>3 195</td>
<td>1 78</td>
<td>2 117</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Norway</td>
<td>3 154</td>
<td>1 110</td>
<td>2 44</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Argentina</td>
<td>3 137</td>
<td>-</td>
<td>-</td>
<td>2 72</td>
<td>1 65</td>
</tr>
<tr>
<td>Switzerland</td>
<td>2 323</td>
<td>-</td>
<td>2 323</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>China</td>
<td>2 280</td>
<td>1 60</td>
<td>-</td>
<td>-</td>
<td>1 220</td>
</tr>
<tr>
<td>New Zealand</td>
<td>2 216</td>
<td>-</td>
<td>2 216</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Japan</td>
<td>2 188</td>
<td>-</td>
<td>2 188</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Brazil</td>
<td>2 98</td>
<td>-</td>
<td>2 98</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Greece</td>
<td>1 295</td>
<td>1 295</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mexico</td>
<td>1 225</td>
<td>1 225</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>1 220</td>
<td>1 220</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Argentina</td>
<td>1 153</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1 153</td>
</tr>
<tr>
<td>Israel</td>
<td>1 150</td>
<td>1 150</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Continues on Next Page with More Studies Separated By Country
<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Civil Psychiatric</th>
<th>Forensic Psychiatric</th>
<th>Correctional</th>
<th>Mixed</th>
<th>Juvenile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Entries</td>
<td>Total N</td>
<td>Entries</td>
<td>Total N</td>
<td>Entries</td>
<td>Total N</td>
</tr>
<tr>
<td>Lithuania</td>
<td>1</td>
<td>118</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>118</td>
</tr>
<tr>
<td>Scotland</td>
<td>1</td>
<td>109</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>109</td>
</tr>
<tr>
<td>Serbia</td>
<td>1</td>
<td>104</td>
<td>1</td>
<td>104</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>France</td>
<td>1</td>
<td>60</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Romania</td>
<td>1</td>
<td>59</td>
<td>1</td>
<td>59</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

3 The number of entries and total N do not always add up to match the totals row due to missing information from some studies.
Table 6. A Summary of Selected HCR-20V1/V2 Research: Key Findings

<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>Means (SD)</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>HCR Total</td>
<td>H Scale</td>
<td>C Scale</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Civil Psychiatric</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alexander et al. (2012)</td>
<td>362</td>
<td>26.00 (5.60)</td>
<td>14.60 (2.90)</td>
<td>6.50 (2.10)</td>
</tr>
<tr>
<td>Arbach-Lucioni et al. (2011)</td>
<td>78</td>
<td>27.10 (5.80)</td>
<td>14.40 (3.00)</td>
<td>7.20 (2.10)</td>
</tr>
<tr>
<td>Blanchard et al. (2013)</td>
<td>139</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Douglas, Ogloff, Nicholls, Grant (1999)</td>
<td>193</td>
<td>19.00 (5.80)</td>
<td>9.80 (3.40)</td>
<td>4.10 (1.90)</td>
</tr>
<tr>
<td>Doyle et al. (2012)</td>
<td>114</td>
<td>15.24</td>
<td>7.97</td>
<td>--</td>
</tr>
<tr>
<td>Garcia-Mansilla (2011)</td>
<td>827</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Gonzales, Nesi, &amp; Yanos (2015)</td>
<td>131</td>
<td>29.00(8.40)</td>
<td>27.10 (8.90)</td>
<td>--</td>
</tr>
</tbody>
</table>

Continues on Next Page with More Civil Psychiatric Samples
<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>Means (SD)</th>
<th></th>
<th></th>
<th></th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>HCR Total</td>
<td>H Scale</td>
<td>C Scale</td>
<td>R Scale</td>
<td>AUCs</td>
<td></td>
</tr>
<tr>
<td>Greig et al. (2013)</td>
<td>106</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AUCs</td>
<td></td>
<td></td>
<td></td>
<td>(.56, .48, .38, .60)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(SPJ, Total, H, C, R)</td>
<td></td>
</tr>
<tr>
<td>Greig et al. (2015)</td>
<td>103</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AUCs</td>
<td></td>
<td></td>
<td></td>
<td>(.58, .64, .60, .63, .59)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Total, H, C, R, SPJ)</td>
<td></td>
</tr>
<tr>
<td>Ho et al. (2015)</td>
<td>220</td>
<td>24.50 (6.80)</td>
<td>12.60 (4.00)</td>
<td>5.50 (2.20)</td>
<td>6.40 (2.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AUCs</td>
<td></td>
<td></td>
<td></td>
<td>(.56, .48, .38, .60)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(SPJ, Total, H, C, R)</td>
<td></td>
</tr>
<tr>
<td>Ivgi et al. (2015)</td>
<td>150</td>
<td>11.33 (5.72)</td>
<td>3.71 (3.38)</td>
<td>4.89 (1.60)</td>
<td>2.72 (2.05)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AUCs</td>
<td></td>
<td></td>
<td></td>
<td>(.56, .48, .38, .60)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(SPJ, Total, H, C, R)</td>
<td></td>
</tr>
<tr>
<td>Jovanović et al. (2009)</td>
<td>104</td>
<td>19.00 (5.30)</td>
<td>17.00 (5.60)</td>
<td>15.00 (4.90)</td>
<td>13.00 (4.20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AUCs</td>
<td></td>
<td></td>
<td></td>
<td>(.56, .48, .38, .60)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(SPJ, Total, H, C, R)</td>
<td></td>
</tr>
<tr>
<td>Martinaki et al. (2013)</td>
<td>295</td>
<td>28.30 (4.40)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AUCs</td>
<td></td>
<td></td>
<td></td>
<td>(.56, .48, .38, .60)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(SPJ, Total, H, C, R)</td>
<td></td>
</tr>
<tr>
<td>McNiel et al. (2003)</td>
<td>100</td>
<td>18.00 (6.60)</td>
<td>7.10 (3.50)</td>
<td>6.10 (2.30)</td>
<td>4.80 (2.30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AUCs</td>
<td></td>
<td></td>
<td></td>
<td>(.56, .48, .38, .60)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(SPJ, Total, H, C, R)</td>
<td></td>
</tr>
<tr>
<td>Morrissey, Beebly, &amp; Milton (2014)</td>
<td>532</td>
<td>23.76 (5.29)</td>
<td>14.70 (3.17)</td>
<td>5.68 (2.63)</td>
<td>3.51 (2.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AUCs</td>
<td></td>
<td></td>
<td></td>
<td>(.56, .48, .38, .60)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(SPJ, Total, H, C, R)</td>
<td></td>
</tr>
<tr>
<td>Nicholls, Ogloff, Douglas (2001)</td>
<td>279</td>
<td>20.40 (5.60)</td>
<td>10.80 (3.30)</td>
<td>7.40 (1.50)</td>
<td>5.40 (2.40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AUCs</td>
<td></td>
<td></td>
<td></td>
<td>(.56, .48, .38, .60)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(SPJ, Total, H, C, R)</td>
<td></td>
</tr>
<tr>
<td>Petris &amp; Podea (2013)</td>
<td>59</td>
<td>19.41 (5.59)</td>
<td>7.05 (3.03)</td>
<td>6.83 (1.66)</td>
<td>5.58 (2.43)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continues on Next Page with More Civil Psychiatric Samples
<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>Means (SD)</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>HCR Total</td>
<td>H Scale</td>
<td>C Scale</td>
</tr>
<tr>
<td>Quin, Li, &amp; Wang (2010)</td>
<td>60</td>
<td>19.10 (6.20)</td>
<td>8.70 (4.00)</td>
<td>5.60 (1.90)</td>
</tr>
<tr>
<td>Ross, Hart, &amp; Webster (1998)</td>
<td>131</td>
<td>17.60 (6.50)</td>
<td>6.50 (3.60)</td>
<td>5.20 (1.70)</td>
</tr>
<tr>
<td>Sada et al. (2016)</td>
<td>225</td>
<td>19.52 (6.33)</td>
<td>24.34 (5.63)</td>
<td>15.17 (5.93)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(SPJ, Total, H, C, R)</td>
<td>(SPJ, Total, H, C, R)</td>
<td>(Total, H, C, R)</td>
</tr>
<tr>
<td>Strub, Douglas, &amp; Nicholls (2016)</td>
<td>100</td>
<td>22.93 (7.29)</td>
<td>12.35 (3.50)</td>
<td>4.91 (2.58)</td>
</tr>
<tr>
<td>Viljoen et al. (2016)</td>
<td>102</td>
<td>20.90</td>
<td>14.40</td>
<td>4.70</td>
</tr>
<tr>
<td>Wilson, Hart, et al. (2009)</td>
<td>131</td>
<td>28.80 (6.20)</td>
<td>13.80 (3.30)</td>
<td>5.50 (2.20)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Total, H, C, R)</td>
<td>(Total, H, C, R)</td>
<td>(Total, H, C, R)</td>
</tr>
<tr>
<td>Allen &amp; Howells (2008)</td>
<td>62</td>
<td>22.70 (6.50)</td>
<td>12.80 (3.60)</td>
<td>4.50 (2.50)</td>
</tr>
<tr>
<td>Arai et al. (2016)</td>
<td>117</td>
<td>14.96 (6.56)</td>
<td>8.42 (3.61)</td>
<td>2.57 (2.16)</td>
</tr>
</tbody>
</table>

Continues on Next Page with More Forensic Samples
<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>HCR Total</th>
<th>H Scale</th>
<th>C Scale</th>
<th>R Scale</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chu et al. (2011)</td>
<td>66</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUC (C) = .68</td>
<td>--</td>
</tr>
<tr>
<td>Chu, Daffern, &amp; Ogloff (2013)</td>
<td>70</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUCs (Total) = .62 - .78</td>
<td>--</td>
</tr>
<tr>
<td>Claix et al. (2002)</td>
<td>86</td>
<td>23.30 (6.30)</td>
<td>12.40 (3.80)</td>
<td>5.00 (2.20)</td>
<td>6.00 (2.00)</td>
<td>r_{tot} = .30 with assault</td>
<td>ICC (Total) = .73</td>
</tr>
<tr>
<td>Crocker &amp; Côté (2009)</td>
<td>60F</td>
<td>21.97 (5.60)</td>
<td>13.75 (3.17)</td>
<td>4.65 (2.65)</td>
<td>3.54 (1.74)</td>
<td>--</td>
<td>ICC (Total) = .87</td>
</tr>
<tr>
<td>de Borba Telles, Folino, &amp; Taborda (2012)</td>
<td>68</td>
<td>23.32 (6.91)</td>
<td>12.35 (3.89)</td>
<td>4.65 (2.37)</td>
<td>6.32 (1.99)</td>
<td>AUCs = .73, .65, .75, .69</td>
<td>(Total, H, C, R)</td>
</tr>
<tr>
<td>Dernevik (1998)</td>
<td>66</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>IRR = .76 - .96</td>
</tr>
<tr>
<td>Dernevik et al. (2001)</td>
<td>826</td>
<td>22.70 (6.50)</td>
<td>12.60 (3.50)</td>
<td>4.80 (1.70)</td>
<td>--</td>
<td>mult R = .66 between HCR &amp; Feeling Checklist</td>
<td>--</td>
</tr>
<tr>
<td>Dernevik et al. (2002)</td>
<td>54</td>
<td>23.10 (5.80)</td>
<td>12.60 (3.50)</td>
<td>4.80 (1.70)</td>
<td>--</td>
<td>AUC = .84 (community)</td>
<td>--</td>
</tr>
<tr>
<td>Desmarais et al. (2012)</td>
<td>120</td>
<td>24.90 (6.59)</td>
<td>13.82 (3.41)</td>
<td>4.81 (2.51)</td>
<td>6.19 (2.36)</td>
<td>AUCs = .80, .73, .74, .77, .79 (Total, H, C, R, SPJ)</td>
<td>ICC (24 Cases) = .71, .88</td>
</tr>
<tr>
<td>de Vogel et al. (2001)</td>
<td>60</td>
<td>26.10 (6.50)</td>
<td>14.60 (3.30)</td>
<td>5.30 (2.20)</td>
<td>6.10 (2.10)</td>
<td>--</td>
<td>IRR (Total) = .79</td>
</tr>
<tr>
<td>de Vogel et al. (2004)</td>
<td>120</td>
<td>22.80 - 32.00</td>
<td>12.60 - 16.00</td>
<td>3.70 - 7.00</td>
<td>5.60 - 9.10</td>
<td>AUC (Total) = .82</td>
<td>IRR (Total) = .83</td>
</tr>
<tr>
<td>de Vries Robbé, de Vogel, &amp; de Spa (2011)</td>
<td>126</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUCs = .68 - .81</td>
<td>--</td>
</tr>
<tr>
<td>de Vries Robbé, de Vogel, &amp; van den Broek (2012)</td>
<td>325</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUC (Total) = .79</td>
<td>--</td>
</tr>
<tr>
<td>de Vries Robbé, de Vogel, &amp; Douglas (2013)</td>
<td>188</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUCs (Total) = .64 - .84</td>
<td>ICC (Total, 24 Cases) = .74</td>
</tr>
<tr>
<td>de Vries Robbé et al. (2014)</td>
<td>83</td>
<td>22.69 (5.79)</td>
<td>13.08 (2.68)</td>
<td>3.76 (2.12)</td>
<td>5.93 (2.40)</td>
<td>AUC = .67 (Total, SPJ)</td>
<td>ICC (Total, 30 Cases) = .86</td>
</tr>
</tbody>
</table>

Continues on Next Page with More Forensic Samples
<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>Means (SD)</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HCR Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Vries Robbé et al. (2015)</td>
<td>108</td>
<td>28.16 (5.00)</td>
<td>13.22 (3.11)</td>
<td>AUCs = .68, .52, .70, .65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23.46 (6.11)</td>
<td>13.48 (2.95)</td>
<td>(Total, H, C, R change scores)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T1</td>
<td>T2</td>
<td>ICCs (19 Cases) = .75, .84, .62, .55 (Total, H, C, R)</td>
</tr>
<tr>
<td>de Vogel et al. (2015)</td>
<td>550</td>
<td>--</td>
<td>--</td>
<td>AUCs = .72, .66, .73, .72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>--</td>
<td>--</td>
<td>(Total, H, C, R)</td>
</tr>
<tr>
<td>Dolan &amp; Fullam (2007)</td>
<td>136</td>
<td>20.50 (6.10)</td>
<td>10.10 (3.50)</td>
<td>AUC (Total) = .67 - .85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T1</td>
<td>T2</td>
<td>--</td>
</tr>
<tr>
<td>Dolan &amp; Khawaja (2004)</td>
<td>70</td>
<td>19.40 (5.70)</td>
<td>11.80 (3.70)</td>
<td>AUC (Total) = .67 - .85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T1</td>
<td>T2</td>
<td>--</td>
</tr>
<tr>
<td>Douglas et al. (1998)</td>
<td>175</td>
<td>24.60 (5.80)</td>
<td>12.50 (3.60)</td>
<td>AUC (Total) = .67 - .70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.50 (2.50)</td>
<td>6.60 (2.30)</td>
<td>Odds ratios = 2.20 - 3.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>--</td>
<td>--</td>
<td>α (Total) = .78</td>
</tr>
<tr>
<td>Douglas et al. (2003)</td>
<td>100</td>
<td>24.70 (4.60)</td>
<td>14.40 (2.80)</td>
<td>AUC (Total) = .67 - .70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.70 (2.00)</td>
<td>5.90 (1.50)</td>
<td>AUC (SPJ) = .68 - .74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>--</td>
<td>--</td>
<td>ICC (Total) = .81</td>
</tr>
<tr>
<td>Fagan et al. (2009)</td>
<td>81</td>
<td>21.10 (9.20)</td>
<td>14.30 (5.80)</td>
<td>AUC (Total) = .76 - .80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.50 (3.30)</td>
<td>2.50 (2.60)</td>
<td>AUC (Total) = .76 - .80</td>
</tr>
<tr>
<td>FitzGerald et al. (2009)</td>
<td>70</td>
<td>26.60 (4.50)</td>
<td>14.92 (2.20)</td>
<td>AUCs = .66 - .88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.28 (2.00)</td>
<td>4.36 (1.80)</td>
<td>AUCs = .66 - .88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>--</td>
<td>--</td>
<td>α (Total) = .78</td>
</tr>
<tr>
<td>Fujii, Lichton et al. (2004)</td>
<td>169</td>
<td>--</td>
<td>--</td>
<td>AUC (Total) = .61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>--</td>
<td>--</td>
<td>AUC (SPJ) = .70</td>
</tr>
<tr>
<td>Fujii et al. (2005)</td>
<td>169</td>
<td>--</td>
<td>--</td>
<td>AUC (SPJ) = .70</td>
</tr>
<tr>
<td></td>
<td></td>
<td>--</td>
<td>--</td>
<td>AUC (AA) = .58</td>
</tr>
<tr>
<td>Grann et al. (2000)</td>
<td>404</td>
<td>--</td>
<td>11.80 (3.70)</td>
<td>AUC (Total) = .61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>--</td>
<td>--</td>
<td>AUC (SPJ) = .70</td>
</tr>
<tr>
<td>Gray et al. (2004)</td>
<td>315</td>
<td>19.90 (7.00)</td>
<td>11.40 (4.00)</td>
<td>AUCs = .56, .48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.80 (2.40)</td>
<td>4.70 (2.60)</td>
<td>AUCs = .56, .48, .62</td>
</tr>
<tr>
<td>Gray et al. (2007)</td>
<td>1,141</td>
<td>22.30 (6.00)</td>
<td>12.40 (3.20)</td>
<td>AUCs = .79, .71, .64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17.90 (6.30)</td>
<td>11.20 (3.80)</td>
<td>AUCs = .68, .55, .63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>N</td>
<td>IRR (Total) = .80 - .88</td>
</tr>
</tbody>
</table>

Continues on Next Page with More Forensic Samples
<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>Means (SD)</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>HCR Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gray, Taylor, &amp; Snowden</td>
<td>887</td>
<td>18.30 (6.70)</td>
<td>11.30 (3.70)</td>
<td>AUCs = .70, .68, .57, .63</td>
</tr>
<tr>
<td></td>
<td>996</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Gray, Taylor, &amp; Snowden</td>
<td>44</td>
<td>19.40 (3.50)</td>
<td>13.20 (3.20)</td>
<td>AUCs = .56, .54, .60</td>
</tr>
<tr>
<td>(2011)</td>
<td></td>
<td></td>
<td>6.10 (2.00)</td>
<td>(HC, C, R)</td>
</tr>
<tr>
<td></td>
<td>44</td>
<td>21.12 (5.70)</td>
<td>12.10 (3.37)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>68</td>
<td>21.12 (5.70)</td>
<td>10.81 (3.38)</td>
<td></td>
</tr>
<tr>
<td>Grevatt et al. (2004)</td>
<td></td>
<td></td>
<td>5.49 (2.93)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5.28 (1.90)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>62</td>
<td>25.20 (7.50)</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Hilterman et al. (2002)</td>
<td>195</td>
<td>24.80</td>
<td>14.10</td>
<td>AUCs = .69, .68, .66, .62, .69</td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>--</td>
<td>--</td>
<td>(Total, H, C, R, SPJ)</td>
</tr>
<tr>
<td>Ireland et al. (2016)</td>
<td>168</td>
<td>23.80</td>
<td>13.47</td>
<td>AUC (SPJ) = .81</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5.40</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>26.93</td>
<td>14.47</td>
<td></td>
</tr>
<tr>
<td></td>
<td>53</td>
<td>23.80 (6.20)</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>McDermott, Edens, et al.</td>
<td>212</td>
<td>12.10 (4.40)</td>
<td>4.40 (0.80)</td>
<td>AUCs = .67, .55, .64, .67</td>
</tr>
<tr>
<td>(2008)</td>
<td></td>
<td></td>
<td>3.00 (1.80)</td>
<td>(Total, H, C, R)</td>
</tr>
<tr>
<td>McDermott, Quanbeck et al.</td>
<td>108</td>
<td>23.80 (6.20)</td>
<td>--</td>
<td>AUCs = .58, .89</td>
</tr>
<tr>
<td>Mokros et al. (2010)</td>
<td></td>
<td>23.80 (6.20)</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td></td>
<td>393</td>
<td>--</td>
<td>--</td>
<td>AUCs = .67, .63, .61, .61</td>
</tr>
<tr>
<td>Morrissey et al. (2007)</td>
<td>73</td>
<td>22.50</td>
<td>--</td>
<td>AUC (Total) = .68, .77</td>
</tr>
</tbody>
</table>

Continues on Next Page with More Forensic Samples
<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>HCR Total</th>
<th>H Scale</th>
<th>C Scale</th>
<th>R Scale</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Müller-Ibsbener &amp; Jockel (1997)</td>
<td>100</td>
<td>--</td>
<td>11.50 (3.60)</td>
<td>5.20 (1.90)</td>
<td>--</td>
<td>M Kappa (H) = .89</td>
<td>M Kappa (C) = .49</td>
</tr>
<tr>
<td>Müller-Ibsbener et al. (1999)</td>
<td>220</td>
<td>24.90 (5.90)</td>
<td>12.00 (3.40)</td>
<td>5.30 (2.20)</td>
<td>7.60 (1.90)</td>
<td>$r_{sr} = .20 \cdot .40$</td>
<td>Kappa (HCR) = .72</td>
</tr>
<tr>
<td>Neal, Miller, &amp; Shealy (2015)</td>
<td>230</td>
<td>18.92 (5.57)</td>
<td>11.23 (3.71)</td>
<td>3.68 (2.50)</td>
<td>4.83 (2.45)</td>
<td>AUCs = .44, .48, .41, .43</td>
<td>(Total, H, C, R)$^56$</td>
</tr>
<tr>
<td>Müller-Ibsbener &amp; Isberner (1997)</td>
<td>100</td>
<td>20.00 (5.30)</td>
<td>11.20 (3.60)</td>
<td>5.10 (2.50)</td>
<td>3.20 (1.20)</td>
<td>AUCs (Total) = .68 -.77</td>
<td>--</td>
</tr>
<tr>
<td>Nishinaka et al. (2016)</td>
<td>71</td>
<td>18.82 (4.12)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>$r = .31 \cdot .46$</td>
<td>--</td>
</tr>
<tr>
<td>Nowak &amp; Nugter (2011)</td>
<td>104</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUC (Total) = .70</td>
<td>--</td>
</tr>
<tr>
<td>Olsson et al. (2013)</td>
<td>267</td>
<td>--</td>
<td>--</td>
<td>5.82 (2.15)</td>
<td>8.50 (2.19)</td>
<td>AUCs (.78), (.75), (.76), (.77)</td>
<td>--</td>
</tr>
<tr>
<td>O’Reilly et al. (2015)$^6$</td>
<td>89</td>
<td>28.20 (8.00)$^V$</td>
<td>15.20 (4.50)$^V$</td>
<td>7.20 (2.78)$^V$</td>
<td>5.80 (2.29)$^V$</td>
<td>AUCs (.78), (.74), (.72), (.69)</td>
<td>--</td>
</tr>
<tr>
<td>Pedersen et al. (2010)</td>
<td>107</td>
<td>24.54 (7.57)</td>
<td>13.59 (3.84)</td>
<td>4.88 (2.56)</td>
<td>6.07 (2.40)</td>
<td>AUCs (.66), (.68), (.62), (.58)</td>
<td>(Total, H, C, R)$^58$</td>
</tr>
<tr>
<td>Pedersen et al. (2012)</td>
<td>81</td>
<td>25.05 (6.86)</td>
<td>13.73 (3.48)</td>
<td>5.42 (2.48)</td>
<td>5.88 (2.18)</td>
<td>AUC (C) = .78$^60$</td>
<td>IRR (5 Cases) = .86, 1.00 (C, R)$^61$</td>
</tr>
<tr>
<td>Petersen et al. (2012)</td>
<td>41</td>
<td>--</td>
<td>12.45$^{59}$</td>
<td>--</td>
<td>--</td>
<td>AUC (Total) = .78$^60$</td>
<td>--</td>
</tr>
<tr>
<td>Pham (2001)</td>
<td>80</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUC (Total) = .67</td>
<td>IRR (Total) = .90</td>
</tr>
<tr>
<td>Philipse (2002)</td>
<td>69</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUC (Modified) = .90$^61$</td>
<td>--</td>
</tr>
<tr>
<td>Pflueger et al. (2015)$^{54}$</td>
<td>259</td>
<td>18.00 (8.00)$^R$</td>
<td>16.00 (11.00)$^N$</td>
<td>--</td>
<td>--</td>
<td>AUC (Total) = .74</td>
<td>--</td>
</tr>
<tr>
<td>Reimann &amp; Nussbaum (2011)</td>
<td>130</td>
<td>24.44 (7.17)</td>
<td>11.90 (3.86)</td>
<td>5.41 (2.59)</td>
<td>7.13 (2.39)</td>
<td>AUC (Total) = .78</td>
<td>--</td>
</tr>
</tbody>
</table>

Continues on Next Page with More Forensic Samples
<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>HCR Total</th>
<th>H Scale</th>
<th>C Scale</th>
<th>R Scale</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ross et al. (2001)</td>
<td>103</td>
<td>20.20 (5.60)</td>
<td>12.70 (3.50)</td>
<td>3.50 (2.10)</td>
<td>4.10 (2.30)</td>
<td>AUC (Total) = .57; .76</td>
<td>--</td>
</tr>
<tr>
<td>Ribeiro, Tully, &amp; Fotiadou (2015)</td>
<td>45</td>
<td>--</td>
<td>--</td>
<td>5.96 (2.42)B</td>
<td>5.56 (2.46)B</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Sánchez-SanSegundo et al. (2014)</td>
<td>88</td>
<td>25.09 (4.70)CD</td>
<td>14.70 (2.30)CD</td>
<td>4.50 (2.20)CD</td>
<td>6.40 (2.60)CD</td>
<td>--</td>
<td>ICC (16 Cases, Total) = .70</td>
</tr>
<tr>
<td>Schaap et al. (2009)</td>
<td>45</td>
<td>24.80 (5.80)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUCs = .41 -.68</td>
<td>ICC (Total) = .98</td>
</tr>
<tr>
<td>Scharin (1999)</td>
<td>49</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>Odds ratio = 9.63</td>
<td>--</td>
</tr>
<tr>
<td>Selenius, Hellström, &amp; Belfrage (2011)</td>
<td>32</td>
<td>26.67 (5.05)D</td>
<td>15.11 (3.41)D</td>
<td>6.06 (1.51)D</td>
<td>5.05 (2.01)D</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Simpson et al. (2015)</td>
<td>232</td>
<td>20.50(6.49)LS</td>
<td>11.40(0.31)LS</td>
<td>4.30(2.46)LS</td>
<td>4.70(2.38)LS</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Snowden, Gray, &amp; Taylor (2010)</td>
<td>1016</td>
<td>18.50 (6.50)</td>
<td>11.30 (3.70)</td>
<td>3.30 (2.50)</td>
<td>3.80 (2.60)</td>
<td>AUCs = .71, .70, .54, .69</td>
<td>ICC (19 Cases, Total) = .80</td>
</tr>
<tr>
<td>Strand &amp; Belfrage (2001)</td>
<td>63</td>
<td>24.80 (7.00)B</td>
<td>12.90 (3.60)B</td>
<td>5.10 (2.60)B</td>
<td>6.70 (2.90)B</td>
<td>No difference between men and women on scales</td>
<td>Kendall’s Tau-b = .67</td>
</tr>
<tr>
<td>Strand et al. (1999)</td>
<td>85</td>
<td>25.50 (7.90)A</td>
<td>13.80 (4.20)A</td>
<td>5.00 (2.50)A</td>
<td>6.70 (2.80)A</td>
<td>AUC (Total) = .80</td>
<td>Kendall’s Tau-b = .69</td>
</tr>
<tr>
<td>Taylor et al (in prep)</td>
<td>212</td>
<td>19.54 (SD)</td>
<td>12.09 (4.43)</td>
<td>4.41 (2.40)</td>
<td>3.04 (1.78)</td>
<td>AUCs = .72, .68, .67, .62</td>
<td>IR = .89, .93, .82</td>
</tr>
<tr>
<td>Telles et al. (2009)</td>
<td>30</td>
<td>23.73(6.25)</td>
<td>13.10 (3.50)</td>
<td>4.83 (1.97)</td>
<td>5.80 (2.09)</td>
<td>--</td>
<td>ICC = .96, .97, .94, .96</td>
</tr>
<tr>
<td>Tengström (2001)</td>
<td>106</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUC (H) = .76</td>
<td>--</td>
</tr>
<tr>
<td>Study / Sample</td>
<td>N</td>
<td>HCR Total</td>
<td>H Scale</td>
<td>C Scale</td>
<td>R Scale</td>
<td>Validity Indices</td>
<td>Reliability Indices</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----</td>
<td>-----------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>-----------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Thomson et al. (2008)</td>
<td>164</td>
<td>--</td>
<td>13.38 (3.43)</td>
<td>--</td>
<td>--</td>
<td>AUC (H) = .53 - .80&lt;sup&gt;69&lt;/sup&gt;</td>
<td>--</td>
</tr>
<tr>
<td>Tiegreen (2010)</td>
<td>54</td>
<td>19.65 (6.46)</td>
<td>12.37 (3.72)</td>
<td>4.91 (2.62)</td>
<td>2.35 (2.18)</td>
<td>AUCs = .67, .58, .64, .66 (Total, H, C, R)</td>
<td>α (Total) = .79</td>
</tr>
<tr>
<td>Urheim et al. (2003)</td>
<td>51</td>
<td>23.50 (6.80)</td>
<td>13.80 (4.30)</td>
<td>5.90 (1.90)</td>
<td>3.90 (2.00)</td>
<td>AUCs = .82, .77, .73, .76&lt;sup&gt;71&lt;/sup&gt;</td>
<td>--</td>
</tr>
<tr>
<td>Vicenzutto, Ducro, &amp; Pham (2016)&lt;sup&gt;55&lt;/sup&gt;</td>
<td>285</td>
<td>23.67 (5.76)&lt;sup&gt;I&lt;/sup&gt;</td>
<td>11.69 (4.04)&lt;sup&gt;I&lt;/sup&gt;</td>
<td>5.19 (1.76)&lt;sup&gt;I&lt;/sup&gt;</td>
<td>6.04 (2.22)&lt;sup&gt;I&lt;/sup&gt;</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Vincent (1998)</td>
<td>125</td>
<td>22.30 (6.30)</td>
<td>11.20 (3.60)</td>
<td>5.00 (2.60)</td>
<td>6.10 (2.10)</td>
<td>Odds ratio = 2.45&lt;sup&gt;72&lt;/sup&gt;</td>
<td>--</td>
</tr>
<tr>
<td>Vitacco et al. (2016)&lt;sup&gt;73&lt;/sup&gt;</td>
<td>116</td>
<td>20.79 (7.57)</td>
<td>12.50 (4.10)</td>
<td>4.55 (3.00)</td>
<td>3.73 (3.26)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Vojt et al. (2009)</td>
<td>115</td>
<td>25.42 (5.33)</td>
<td>15.46 (2.87)</td>
<td>5.22 (2.43)</td>
<td>4.73 (2.50)</td>
<td>AUCs = .44, .42, .55, .46 (Total, H, C, R)</td>
<td>--</td>
</tr>
<tr>
<td>Vojt, Thomson, &amp; Marshall (2013)&lt;sup&gt;74&lt;/sup&gt;</td>
<td>109</td>
<td>25.22 (5.10)</td>
<td>15.30 (2.70)</td>
<td>5.22 (2.40)</td>
<td>4.70 (2.40)</td>
<td>AUCs = .86, .68, .79, .75 (Total, H, C, R)</td>
<td>--</td>
</tr>
<tr>
<td>Whittemore (1999)</td>
<td>172</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>Wald = 9.86</td>
<td>--</td>
</tr>
<tr>
<td>Wilson, Desmarais et al. (2009)&lt;sup&gt;75&lt;/sup&gt;</td>
<td>30</td>
<td>24.97 (6.58)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUCs = .72 - .86</td>
<td>--</td>
</tr>
<tr>
<td>Wintrup (1996)&lt;sup&gt;76&lt;/sup&gt;</td>
<td>80</td>
<td>--&lt;sup&gt;76&lt;/sup&gt;</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>Mean r = .30</td>
<td>--</td>
</tr>
</tbody>
</table>

**Correctional**

<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>HCR Total</th>
<th>H Scale</th>
<th>C Scale</th>
<th>R Scale</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basdeskis-Jozsa et al. (2013)</td>
<td>53</td>
<td>25.84 (6.23)&lt;sup&gt;77&lt;/sup&gt;</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Belfrage, Fransson, &amp; Strand (2000)&lt;sup&gt;78&lt;/sup&gt;</td>
<td>41</td>
<td>26.00</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>Cohen’s d = 1.70, 1.00, 1.14, 1.22 (Total, H, C, R)</td>
<td>--</td>
</tr>
<tr>
<td>Blais (2014)&lt;sup&gt;79&lt;/sup&gt;</td>
<td>112</td>
<td>28.60 (4.60)&lt;sup&gt;PA&lt;/sup&gt;</td>
<td>23.80 (8.20)&lt;sup&gt;CA&lt;/sup&gt;</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Continues on Next Page with More Correctional Samples
<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>Means (SD)</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>HCR Total</td>
<td>H Scale</td>
<td>C Scale</td>
</tr>
<tr>
<td>Coid et al. (2009)</td>
<td>304B</td>
<td>20.20 (7.48)B</td>
<td>12.10 (4.29)B</td>
<td>3.11 (2.11)B</td>
</tr>
<tr>
<td></td>
<td>1353A</td>
<td>19.10 (7.80)A</td>
<td>11.10 (4.56)A</td>
<td>3.39 (2.15)A</td>
</tr>
<tr>
<td>Coid et al. (2015)</td>
<td>1,001</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Cooke et al. (2001)</td>
<td>250</td>
<td>--</td>
<td>10.90</td>
<td>--</td>
</tr>
<tr>
<td>Cooke et al. (2015)</td>
<td>22C</td>
<td>--</td>
<td>11.70 (4.30)C</td>
<td>3.90 (2.20)C</td>
</tr>
<tr>
<td>Côté (2001)</td>
<td>36F</td>
<td>--</td>
<td>13.10 (3.50)F</td>
<td>4.80 (2.50)F</td>
</tr>
<tr>
<td>Couplan &amp; Olver (2015)</td>
<td>178</td>
<td>28.30 (5.50)T1</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Dahle (2002)</td>
<td>200</td>
<td>19.00 (6.50)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Douglas &amp; Belfrage (2002)</td>
<td>72</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Douglas &amp; Webster (1999)</td>
<td>188</td>
<td>20.10 (7.90)</td>
<td>11.10 (3.90)</td>
<td>4.40 (2.50)</td>
</tr>
<tr>
<td>Doyle &amp; Dolan (2003)</td>
<td>129</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Doyle et al. (2002)</td>
<td>87</td>
<td>--</td>
<td>14.10 (3.90)R</td>
<td>11.40 (3.90)N</td>
</tr>
<tr>
<td>Dunbar (2003)</td>
<td>58</td>
<td>22.20 (9.90)</td>
<td>9.40 (5.70)</td>
<td>6.10 (1.90)</td>
</tr>
<tr>
<td>Folino (2015)</td>
<td>153</td>
<td>20.10 (6.90)</td>
<td>11.70 (4.20)</td>
<td>4.70 (1.90)</td>
</tr>
</tbody>
</table>

Continues on Next Page with More Correctional Samples
<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>Means (SD)</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>HCR Total</td>
<td>H Scale</td>
<td>C Scale</td>
</tr>
<tr>
<td>Freese et al. (2002)</td>
<td>128</td>
<td>22.20 (6.50)&lt;sup&gt;R&lt;/sup&gt;</td>
<td>12.80 (3.70)&lt;sup&gt;R&lt;/sup&gt;</td>
<td>3.90 (2.80)&lt;sup&gt;R&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17.20 (6.50)&lt;sup&gt;N&lt;/sup&gt;</td>
<td>10.40 (4.40)&lt;sup&gt;N&lt;/sup&gt;</td>
<td>3.10 (1.90)&lt;sup&gt;N&lt;/sup&gt;</td>
</tr>
<tr>
<td>Gray et al. (2003)</td>
<td>34</td>
<td>--</td>
<td>10.20 (3.40)</td>
<td>5.40 (2.60)&lt;sup&gt;N&lt;/sup&gt;</td>
</tr>
<tr>
<td>Guy et al. (2015)&lt;sup&gt;86&lt;/sup&gt;</td>
<td>5187</td>
<td>11.78 (4.44)&lt;sup&gt;P&lt;/sup&gt;</td>
<td>8.10 (3.41)&lt;sup&gt;P&lt;/sup&gt;</td>
<td>0.76 (1.15)&lt;sup&gt;P&lt;/sup&gt;</td>
</tr>
<tr>
<td>Howard (2007)</td>
<td>64</td>
<td>10.30 (3.40)&lt;sup&gt;88&lt;/sup&gt;</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17.80 (8.30)&lt;sup&gt;88&lt;/sup&gt;</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Kroner &amp; Mills (2001)</td>
<td>97</td>
<td>17.80 (8.30)&lt;sup&gt;88&lt;/sup&gt;</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Kroner &amp; Mills (2011)</td>
<td>248</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Liu et al. (2011)&lt;sup&gt;90&lt;/sup&gt;</td>
<td>1125</td>
<td>22.21 (6.55)&lt;sup&gt;VR&lt;/sup&gt;</td>
<td>13.07(3.88)&lt;sup&gt;VR&lt;/sup&gt;</td>
<td>4.08 (2.04)&lt;sup&gt;VR&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17.88 (7.84)&lt;sup&gt;N&lt;/sup&gt;</td>
<td>10.37(4.69)&lt;sup&gt;N&lt;/sup&gt;</td>
<td>3.15 (2.11)&lt;sup&gt;N&lt;/sup&gt;</td>
</tr>
<tr>
<td>Livingston et al. (2015)</td>
<td>250</td>
<td>17.77 (5.48)&lt;sup&gt;N&lt;/sup&gt;</td>
<td>11.56 (3.59)</td>
<td>2.97 (2.19)</td>
</tr>
<tr>
<td>Mills et al. (2007)</td>
<td>83</td>
<td>18.30 (8.40)&lt;sup&gt;48&lt;/sup&gt;</td>
<td>9.20 (4.40)</td>
<td>4.10 (2.20)</td>
</tr>
<tr>
<td>Mills &amp; Gray (2013)</td>
<td>78</td>
<td>19.28 (7.67)&lt;sup&gt;N&lt;/sup&gt;</td>
<td>9.68 (4.20)</td>
<td>4.32 (2.00)</td>
</tr>
<tr>
<td>Neves, Goncalves, &amp; Palma - Oliveira (2011)</td>
<td>158</td>
<td>15.34 (6.64)&lt;sup&gt;N&lt;/sup&gt;</td>
<td>7.52 (3.70)</td>
<td>3.61 (2.00)</td>
</tr>
<tr>
<td>Pham et al. (2000)</td>
<td>68</td>
<td>20.70 (9.20)&lt;sup&gt;N&lt;/sup&gt;</td>
<td>10.40 (5.00)</td>
<td>4.80 (2.40)</td>
</tr>
<tr>
<td>Stadtland (2008)</td>
<td>86</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Continues on Next Page with More Correctional Samples
<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>HCR Total</th>
<th>H Scale</th>
<th>C Scale</th>
<th>R Scale</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sturup et al. (2015)</td>
<td>98</td>
<td>15.80 (8.30)</td>
<td>9.50 (4.70)</td>
<td>2.40 (2.20)</td>
<td>3.90 (2.90)</td>
<td>AUC = .66, .62, .72, .47 (Total, H, C, R)</td>
<td>--</td>
</tr>
<tr>
<td>Verbrugge, Goodman-Delahunt, &amp; Frize (2011)</td>
<td>59</td>
<td>26.93 (5.13)</td>
<td>14.69 (3.01)</td>
<td>6.15 (1.74)</td>
<td>6.19 (1.75)</td>
<td>AUCs = .80,.75,.67,.75,.81 (Total, H, C, R, SPJ)</td>
<td>ICC (7 Cases) = .67</td>
</tr>
<tr>
<td>Vincent (1998)</td>
<td>125</td>
<td>23.60 (6.70)</td>
<td>11.90 (3.80)</td>
<td>4.50 (2.50)</td>
<td>7.30 (1.70)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Vincent et al. (2001)</td>
<td>56</td>
<td>--</td>
<td>--</td>
<td>3.90 (2.60)\textsuperscript{RE}</td>
<td>5.80 (2.10)\textsuperscript{RE}</td>
<td>--</td>
<td>ICC\textsubscript{1} = .70, .58 (C, R)</td>
</tr>
<tr>
<td>Wellbeloved-Stone et al. (2016)</td>
<td>471</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUCs = .73,.71,.68,.69 (Total, H, C, R)\textsuperscript{A}</td>
<td>--</td>
</tr>
<tr>
<td>Blanchard &amp; Douglas (2014)</td>
<td>235</td>
<td>--</td>
<td>--</td>
<td>4.44 (2.31)</td>
<td>5.08 (2.40)</td>
<td>Odds ratios = 1.25, 1.21 (C, R)</td>
<td>ICC = .63, .42 (C, R)</td>
</tr>
<tr>
<td>Cornelis, Joyal, &amp; Côté (2011)</td>
<td>178</td>
<td>20.70 (5.60)\textsuperscript{FN}</td>
<td>18.10 (7.10)\textsuperscript{FN}</td>
<td>28.40 (5.00)\textsuperscript{CV}</td>
<td>25.50 (4.30)\textsuperscript{CN}</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Hill et al. (2012)</td>
<td>90</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUCs = .66,.62,.69,.43 (HC, H, C, R)\textsuperscript{98}</td>
<td>--</td>
</tr>
<tr>
<td>Ho et al. (2009)</td>
<td>96</td>
<td>--</td>
<td>13.14 (4.42)</td>
<td>--</td>
<td>--</td>
<td>AUCs (H) = .61 - .74</td>
<td>--</td>
</tr>
<tr>
<td>Ho et al. (2013)</td>
<td>220</td>
<td>16.59 (6.80)</td>
<td>8.37 (3.98)</td>
<td>4.25 (2.30)</td>
<td>3.98 (2.19)</td>
<td>AUCs = .68 -.70\textsuperscript{99}</td>
<td>ICC (110 Cases) = .57,.71,.43,.37,.73 (Total, H, C, R, SPJ)</td>
</tr>
<tr>
<td>Hodgins et al. (2001)</td>
<td>126</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>ICC\textsubscript{1} = .85-.99 between raters and criteria scores</td>
<td>ICC\textsubscript{1} = .90,.94,.89,.68 (Total, H, C, R)</td>
</tr>
</tbody>
</table>

**Mixed Samples**

<table>
<thead>
<tr>
<th>Study / Sample</th>
<th>N</th>
<th>HCR Total</th>
<th>H Scale</th>
<th>C Scale</th>
<th>R Scale</th>
<th>Validity Indices</th>
<th>Reliability Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blanchard &amp; Douglas (2014)</td>
<td>235</td>
<td>--</td>
<td>--</td>
<td>4.44 (2.31)</td>
<td>5.08 (2.40)</td>
<td>Odds ratios = 1.25, 1.21 (C, R)</td>
<td>ICC = .63, .42 (C, R)</td>
</tr>
<tr>
<td>Cornelis, Joyal, &amp; Côté (2011)</td>
<td>178</td>
<td>20.70 (5.60)\textsuperscript{FN}</td>
<td>18.10 (7.10)\textsuperscript{FN}</td>
<td>28.40 (5.00)\textsuperscript{CV}</td>
<td>25.50 (4.30)\textsuperscript{CN}</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Hill et al. (2012)</td>
<td>90</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>AUCs = .66,.62,.69,.43 (HC, H, C, R)\textsuperscript{98}</td>
<td>--</td>
</tr>
<tr>
<td>Ho et al. (2009)</td>
<td>96</td>
<td>--</td>
<td>13.14 (4.42)</td>
<td>--</td>
<td>--</td>
<td>AUCs (H) = .61 - .74</td>
<td>--</td>
</tr>
<tr>
<td>Ho et al. (2013)</td>
<td>220</td>
<td>16.59 (6.80)</td>
<td>8.37 (3.98)</td>
<td>4.25 (2.30)</td>
<td>3.98 (2.19)</td>
<td>AUCs = .68 -.70\textsuperscript{99}</td>
<td>ICC (110 Cases) = .57,.71,.43,.37,.73 (Total, H, C, R, SPJ)</td>
</tr>
<tr>
<td>Hodgins et al. (2001)</td>
<td>126</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>ICC\textsubscript{1} = .85-.99 between raters and criteria scores</td>
<td>ICC\textsubscript{1} = .90,.94,.89,.68 (Total, H, C, R)</td>
</tr>
<tr>
<td>Study / Sample</td>
<td>N</td>
<td>Means (SD)</td>
<td>H Scale</td>
<td>C Scale</td>
<td>R Scale</td>
<td>Validity Indices</td>
<td>Reliability Indices</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>----</td>
<td>---------------------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------</td>
<td>------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HCR Total</td>
<td>H Scale</td>
<td>C Scale</td>
<td>R Scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michel et al. (2013)</td>
<td>248</td>
<td>18.71 (6.83)</td>
<td>11.65 (3.91)</td>
<td>3.18 (2.11)</td>
<td>3.87 (2.35)</td>
<td>AUCs = .67 - .74</td>
<td>ICC (35 Cases) = .895, .78, .52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19.19 (7.04)</td>
<td>9.19 (4.50)</td>
<td>4.70 (1.88)</td>
<td>5.30 (2.07)</td>
<td>AUCs = .60 - .74</td>
<td>(H, C, R)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(G)</td>
<td>(G)</td>
<td>(G)</td>
<td>(G)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagi et al. (2009)</td>
<td>49</td>
<td>--</td>
<td>10.68 (3.28)</td>
<td>--</td>
<td>--</td>
<td>AUCs (H) = .52 - .57</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td>--</td>
<td>9.25 (2.51)</td>
<td>--</td>
<td>--</td>
<td>AUCs (H) = .66 - .82</td>
<td>--</td>
</tr>
<tr>
<td>Nanayakkara, O’Driscoll, &amp; Allnutt (2012)</td>
<td>234</td>
<td>25.21 C, O105</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>16.85 F</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Nilsson et al. (2011)</td>
<td>100</td>
<td>--</td>
<td>13.69 (5.61)</td>
<td>6.77 (2.36)</td>
<td>6.60 (2.52)</td>
<td>AUC (Total) = .71</td>
<td>--</td>
</tr>
<tr>
<td>O’Shea et al. (2014)</td>
<td>505</td>
<td>27.05 (5.61)</td>
<td>13.69 (5.61)</td>
<td>6.77 (2.36)</td>
<td>6.60 (2.52)</td>
<td>AUCs = .66, .47, .70, .68, .60</td>
<td>--</td>
</tr>
<tr>
<td>O’Shea et al. (2015)</td>
<td>613</td>
<td>28.73 (5.13)</td>
<td>14.25 (2.66)</td>
<td>7.03 (2.18)</td>
<td>7.49 (2.16)</td>
<td>AUCs = .61, .48, .58, .66, .62</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td>27.04 (5.61) C</td>
<td>13.69 (3.80)</td>
<td>6.77 (2.36)</td>
<td>6.59 (2.52)</td>
<td>AUCs = .61, .44, .66, .65, .58</td>
<td>--</td>
</tr>
<tr>
<td>Pham (2015)</td>
<td>235</td>
<td>22.69</td>
<td>11.90</td>
<td>5.06</td>
<td>5.74</td>
<td>AUC (H) = .75</td>
<td>--</td>
</tr>
<tr>
<td>Spehr et al. (2010)</td>
<td>166</td>
<td>16.91 (4.32) CV</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16.15 (4.82) AV</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

**Juvenile Samples**

MacEachern (2001) | 108  | 19.70 (6.60) | 7.90 (3.20) | 5.60 (1.90) | 6.20 (2.30) | r = .35 - .46 | ICC1 = .86, .88, .80, .77 |
|                  |      | AUCs = .73 - .79 |             |             |             | (Total, H, C, R) |
Note 1. This table does not contain all studies reported in the annotated bibliography. Some studies supplemental to main studies were not included. Other studies were excluded from the table if they addressed issues other than the relationship between the HCR-20 V1/V2 and violence. The method and results of the studies in this table are described in more detail in the annotated bibliography that follows.

Note 2. IRR = interrater reliability; HC = total H Scale and C Scale composite when R Scale not available; SPJ = structured professional judgment of low, moderate, or high risk; Superscript “A” denotes analyses for men only; Superscript “B” denotes analyses for women only. “ψ” denotes that the sample is a sub-sample of another study, and hence the reporting of Ms and SDs is omitted.

1 Superscript “IP” denotes discharged patients with an intellectual disability (ID) and personality disorder (PD), “I” denotes patients with ID only, and “P” denotes patients with PD only.
2 Study provides AUCs for 4, 8, and 12 months. Psychopathy Item (H7) is omitted from the AUC values reported for HCR-20 Total.
3 Superscript “S” denotes short-stay civil psychiatric patients and “L” denotes long-stay civil psychiatric patients.
4 AUC values were provided for several subgroups; range of AUC values is presented.
5 See also Douglas, Ogloff, & Nicholls (1997a, b).
6 Superscript “V” denotes discharged patients who were violent in the community and “N” denotes discharged patients who were not violent.
7 Prior Supervision Failure (H10) could not be coded and was excluded from analyses.
8 AUC values were also reported collapsed across gender.
9 Superscript “AOT” denotes patients on assisted outpatient treatment (AOT) and “C” denotes non-AOT patients.
10 AUCs reported are for physical violence. The authors also present AUCs for any violence and verbal violence.
11 AUCs reported are for any violence at the 12-month follow-up. The authors also report AUCs at the 6-month follow-up and AUCs at the 6- and 12-month follow-ups separately by type and severity of violence.
12 AUCs reported are for any violence at the 12-month follow-up. The authors also reported AUCS at the 6-month follow-up and AUCs at the 6- and 12-month follow-ups separately by type of violence.
13 The AUCs reported here refer to any violence perpetration, the authors also report AUCs for non-physical and physical violence.
14 Superscript “MI” denotes patients in a large mental health service, “PD” denotes patients in a mainstream personality disorder service, “DS” denotes patients in a ‘Dangerous and Severe Personality Disorder’ service, “LD” denotes males patients in a learning disability service, and “B” notes female patients in a learning disability service.
15 See also Nicholls, Ogloff, & Douglas (1997a, b). Table reports validity indices for community violence only.
16 See also Klassen (1996).
17 AUCs reported are for physical violence. The authors also present AUCs for any violence and verbal violence.
18 AUCs reported are for any violence at the 12-month follow-up. The authors also report AUCs separately by type of violence (i.e., physical, verbal, severe) and gender at 6- and 12-month follow-ups.
19 AUC is for any violence but AUCs are available for Level 1 and Level 2 violence.
20 AUCs reported are for the 6-month follow-up. The authors also report AUCs for the 3-month follow-up.
21 AUCs reported are for physical aggression at 180-day follow-up. The authors also reported AUCs separately for any and verbal aggression at the 90- and 180-day follow-ups.
22 AUC reported is for any inpatient aggression. AUCs are also presented for interpersonal violence and verbal threat.
23 Study provides AUCs at 1, 3, and 6 months. AUCs provided are for interpersonal aggression. The authors also present AUCs for verbal threat and any aggression.
Superscript “F” denotes individuals sentenced to a forensic psychiatric hospital, while superscript “C” denotes individuals sentenced to a civil psychiatric hospital following a NCRMD decision.

AUCs were also reported for any antisocial event.

Eight patients were rated by 40 clinicians, where each patient was rated by five different clinicians.

The top mean (22.7) was derived from psychologists; the bottom mean (26.3) from psychiatric nurses.

AUCs reported are for any aggression. The authors also present AUCs separately for verbal aggression, physical aggression towards objects, and physical aggression towards individuals.

Means were provided for four different subgroups; range of means is presented.

Study provides AUCs for 1, 2, and 3 years post-discharge.

AUCs are also reported for the combined HCR-SAPROF total score and final risk judgment.

AUCs were provided by the authors for violent recidivism at 1 year, 3 years, and long term.

AUCs reported are for long-term (i.e., average 15-year) follow-up. The authors also report AUCs for 1- and 3-year follow-ups.

Superscript “T1” denotes pre-treatment scores and “T2” denotes post-treatment scores.

80 cases for follow-up.

Psychopathy Item (H7) is omitted from mean values reported for HCR-20 Total and H-scale.

AUCs were provided for readmission, self/collateral reports of violence, self/collateral reports of re-offending, and serious re-offending; range of AUCs is reported.

Sample overlaps with Douglas et al. (1998); Ross et al (2001); Douglas & Ogloff (2003).

The means reported are taken from Ijaz et al. (2009), which reported descriptive statistics from the same sample.

Superscript “L” denotes the learning disabled group; superscript “C” denotes the control group of psychiatric patients (without a diagnosis of LD).

Validity indices reported are for inpatient violence.


These samples are a combination of forensic and correctional.

Superscript “I” denotes the intellectually disabled group; superscript “N” denotes the non-intellectually disabled group.

The reported AUC is for violent convictions. Study provides AUCs for .5, 1, 2 years follow-up for both violence and any conviction.

AUC values reported are for violent recidivism. The authors also provide AUC values for any recidivism. AUC values for both violent and any recidivism are also reported separately for each of the psychiatric diagnostic categories with a sample size larger than 100.

Total Score is for HC composite.

AUC values reported are for serious (violent and sexual recidivism). The authors also provide AUCs for general recidivism.

AUCs reported are for any aggression at 12-months and 4-years. AUCs are also reported for various categories of aggression towards others and self-harm at 12-month and 4-year follow-ups.

Baseline means. The authors also report means at discharge.

AUC reported is for violent recidivism. The authors also report AUCs for general recidivism, and both outcomes before and after treatment.

The reported IRR is percent agreement.

The reported AUC is for total aggression. AUCs are also available for staff and patient directed aggression.
Superscript “R” denotes individuals that recidivated in the follow-up period; superscript “N” denotes non-recidivists. The first AUC is for interpersonal physical violence and the second is verbal/property violence. AUCs reported are for contact violence. The authors also present AUCs for combined contact and threatened violence, and hospital readmission. This sample also comprises the analyses for Vincent’s (1999) forensic sample. AUCs reported are for violent reconvictions. The authors also present AUCs for inpatient aggression. Baseline mean. AUC reported is for C score at Follow-Up 1 predicting violence at Follow-Up 2. Based on a subset of most predictive items. Ross et al. (2001) overlaps in sample with Douglas et al. (1998). Superscript “CD” denotes patients with severe mental illness and conduct disorder and “N” denotes the comparison group (i.e., patients with severe mental illness and no conduct disorder). These samples are a combination of forensic and correctional. Superscript “D” denotes patients with dyslexia and “N” denotes patients without dyslexia. Superscript “ES” denotes patients for whom the age at which they first experienced psychiatric symptoms was greater than the age when they were first arrested (i.e., early starters), superscript “LS” denotes patients who had suffered fewer than 10 years of mental illness before their first arrest and who were younger than 37.5 years at the time of their first offence (i.e., late starters), and superscript “LLS” denotes patients who had experienced 10 or more years of illness duration before their first arrest and who were 37.5 years or older when they were first arrested (i.e., late late starters). Means and AUC values were also reported separately for black and white participants. The reported IRR is percentage agreement. AUC values represent a range of incidents and violent offences. Authors present AUCs for any incident, any offence, frequency of incidents and frequency of serious incidents. AUC values reported for any violence, but are also available for Category 1 and Category 2 violence. AUC values reported are for most severe episode of inpatient violence. The authors also presented AUCs for frequency. The effect size was for predicting time institutionalized, not violence. Means reported are for the total sample. The authors also present means separately for patients who continued to meet inpatient commitment criteria, patients who were unable to maintain conditional release, and patients who were able to maintain conditional release. AUCs reported are for serious (violent or sexual) incidents. AUCs are also provided for all incidents, minor incidents, and any conviction. See also Douglas, Webster, & Wintrup (1996). The mean for the HC composite was 17.1 (SD = 3.8). Mean presented is for entire sample. Total and subscale means are also presented separately for violent offenders, sex offenders, and robbers. Superscript “PA” denotes HCR-20 scores from prosecution-appointed evaluators and superscript “CA” denotes HCR-20 scores from court-appointed evaluators. AUC values represent violent reconvictions, acquired reconvictions, and any reconvictions. The authors also present AUCs for each scale separated by gender. AUCs reported are for combined official and self-report records of violent convictions. AUCs are also reported separately for official violent convictions and self-reported violent convictions, and all outcomes when non-violent offenders are excluded from the analysis. Reported here for community violence only.
82 Superscript “C” denotes involuntarily civil psychiatric patients, “F” denotes forensic patients (found not criminally responsible on account of mental disorder), and “M” denotes mentally disordered offenders.
83 These analyses are based on re-analysis of existing data sets across three samples, and hence N, descriptive characteristics, and reliability co-efficients are not provided here. The three d score ranges in the validity indices column refer to changes in C and R scores over time and not to any relationship with violence.
84 See also Douglas, Webster, & Wintrup (1996).
85 AUCs reported are for violent recidivism. The authors also provide AUCs for general recidivism.
86 Superscript “P” denotes inmates that were granted parole and “N” denotes inmates not granted parole.
87 AUCs reported are for parole suitability decisions.
88 CR composite.
89 AUCs reported are for violent recidivism but AUCs are available for general recidivism.
90 AUCs are also provided for non-violent recidivism.
91 AUCs presented are for four different validation samples. AUCs are also presented separately for LR, CART, and NN models.
92 AUCs report are for violent behavior. The authors also provide AUCS for non-compliance, a psychiatric adverse event, and any criminal justice contact.
93 AUCs reported are for violent recidivism. The authors also present AUCs for any (i.e., violent and non-violent) recidivism.
94 AUCs reported are for violent recidivism. The authors also present AUCs for general and nonviolent recidivism.
95 Means with superscript “RE” refer to researcher-based ratings; those with superscript “CL” refer to clinician-based ratings.
96 AUCs reported are for combined official and self-reported records of institutional violence. The authors also report AUCs separately for official violence and self-reported violence.
97 Superscript “F” denote participants from forensic settings and “C” denotes participants from correctional settings. Additional superscripts of “V” and “N” denote participants that were violent and not violent, respectively.
98 AUCs reported are for violent recidivism. The authors also present AUCs for general, sexual and violent sexual recidivism.
99 AUCs reported are for any violence. Study provides AUCs for any violence, verbal violence, violence against others and violent convictions at 6 and 12 months. AUCs are also available for subscales and SPJ ratings.
100 There are four subsamples across different countries, with means broken down accordingly. They are not reported because of space.
101 Superscript “F” denotes forensic psychiatric patients and “G” denotes general psychiatric patients.
102 AUCs were reported at 6, 12, 18, and 24 months. AUCs are also available for subscale and final risk ratings.
103 Means with superscript “I” refer to the inpatient sample; those with superscript “O” refer to the community or outpatient sample.
104 Superscript “F” denotes forensic psychiatric patients and “C” denotes civil psychiatric patients.
105 Means were also provided for forensic and civil groups by recruitment location.
106 Means provided are for the total sample. The authors also provided means for patients who and did not engage in any or physical aggression. Psychopathy Item (H7) is omitted from the HCR-20 Total and H means.
107 AUCs are also reported separately by gender, ethnicity, and mental health diagnosis. Psychopathy Item (H7) is omitted from the AUC values reported for HCR-20 Total and H scales.
108 AUCs reported are also reported are for physical aggression. AUCs are also reported separately for ID and comparison groups for any aggression. Psychopathy Item (H7) is omitted from the AUC values reported for HCR-20 Total and H scales.
Means reported are for total sample. The authors also report means separately for patients who were and were not detained under Article 21.

Superscript “CV” denotes sexual murderers with child victims and “AV” denotes sexual murderers with adult victims.
Few studies have systematically examined how patients with both intellectual disability and personality disorders (the ID-PD group) differ from either those with a intellectual disability alone (the ID group) or those with a personality disorder alone (the PD group). The aim of this study was to compare these three groups on a number of pre- and post-treatment variables to establish whether the ID-PD group was more similar to the PD group or to the ID group. As part of this examination group differences on pre-treatment PCL: SV and HCR-20 scores were examined. Post-treatment variables included: post-release convictions, rates of re-offending at 1, 2 and 5 years, and serious/violent reoffending (violent offenses included all those classified as violence against persons by the UK Home Office, as well as kidnap, criminal damage endangering life, Robbephy, rape, and indecent assault) at two years.

This study was a retrospective case analysis of 362 patients who were discharged from medium secure facilities in the UK. Of the sample, 48 subjects had both an intellectual disability and a personality disorder, 97 had an intellectual disability only, and 217 had a personality disorder only. The HCR-20 and PCL: SV were completed blind to any outcome following discharge by research psychologists using file-based information that was available at the time of the patient’s discharge from the unit.

On the HCR-20 there was strong evidence of an overall difference between the three groups on the H, C, R and total scores. For the ID-PD group, mean scores were 14.6 (SD = 2.9), 6.5 (SD = 2.1), 4.9 (SD = 2.2), and 26.0 (SD = 5.6) on the HCR-20 historical, clinical, risk, and total scales, respectively. For the ID group, mean scores were 11.3 (SD = 2.7), 5.5 (SD = 2.0), 3.7 (SD = 2.5), and 20.4 (SD = 5.3) on the HCR-20 historical, clinical, risk and total scores respectively. For the PD group, mean scores were 13.8 (SD = 3.7), 3.5 (SD = 2.3), 5.0 (SD = 2.6), and 22.2 (SD = 6.1) on the HCR-20 historical, clinical, risk and total scores respectively. Pairwise comparisons revealed that the ID-PD group had significantly higher scores on all HCR-20 scales compared to the ID group. Significant differences between the ID-PD group and the PD group were only present on the C scale. Total PCL: SV scores were 5.5 (SD = 2.8), 3.7 (SD = 3.1), 4.8 (SD = 3.7) for the ID-PD, ID and PD groups, respectively. The ID-PD group had significantly higher scores on the PCL: SV compared to the ID group, but there were no significant differences in psychopathy screening scores between the ID-PD group and PD groups.

Convictions were highest in the PD group (37%) compared to the ID-PD (23%) and ID groups (14%). Reoffenses at one, two and five years and a serious/violent reconviction were also more prevalent in the PD group and least prevalent in the ID group. With regards to serious/violent reoffense rates at 2 years, 23% of the PD group compared to 10% of the ID-PD and 3% of the ID groups reoffended.

The authors concluded that higher scores in the ID-PD group on the HCR-20 and PCL: SV were likely due to the presence of the dual diagnosis. In contrast to the two comparison groups, the ID-PD group had more clinical issues which would be relevant to the PCL: SV and HCR-20. In terms of future offending and convictions, it seemed that the ID-PD group appeared more similar to the ID group, however the authors note that these findings should be interpreted with caution. These results may be due to the fact that those with intellectual disabilities are diverted away from the criminal justice system and not subject to the same legal processes as those in the PD group. Further, because those with intellectual disability are more difficult to place they may be subject to longer restrictions than the PD group.
Project and Scholarly Work


Summary

This study evaluated the accuracy of the HCR-20 and PCL: SV for predicting inpatient violent behavior (any incident of physical violence/aggression towards another person or property) over a one-year period in a sample of chronic adult psychiatric inpatients in Spain (N = 78). This was the first prospective research study of the HCR-20’s predictive validity in a Spanish civil mental health hospital. Study participants were selected on the basis of the following criteria: past violence and admission to a long-stay unit. Participants had a mean age of 42.8 years (SD = 9.7). Most participants in the sample were male (74.4%), unemployed (97.4%), and single (75.6%), with zero or low education level (73%). Primary diagnoses included paranoid schizophrenia (55%), other schizophrenia (14.1%), schizoaffective disorder (10.3%), personality disorder (10.3%), organic disorder (3.8%), and other disorders (1.3%). At the time of the study the sample’s mean duration of hospitalization was 1472 days (SD = 1443).

The HCR-20 and PCL: SV were coded by a clinical psychologist using admission summaries, psychiatric assessment reports, and nurse reports within the three months prior to the beginning of the follow-up period. Patients were followed-up while in the hospital for 12 months by nursing staff who recorded violent incidents using a Spanish adaptation of the Modified Overt Aggression Scale. Nursing staff were blind to the scores on the risk assessment tools to avoid bias in data collection. For the purpose of this study only two categories of aggression were considered: aggression against property (wanton and reckless destruction of ward paraphernalia or other possessions) and physical aggression against others (violent action intended to inflict pain, bodily harm, or death upon another). During the one year follow-up period, 53.8% of the sample was involved in at least one physically violent incident against another person and 35.9% were violent towards objects.

Total scores on the HCR-20 (with item 7 removed) were strongly correlated with total scores on the PCL: SV (r = 0.60, p < .01). The average PCL-R total score in the sample was 8.7 (SD = 6.2). The interpersonal (F1) and social deviance (F2) mean scores were 4 (SD = 3.2) and 4.7 (SD = 3.6), respectively. The average HCR-20 total score was 27.1 (SD = 5.8). The subscales’ mean scores were 14.4 (SD = 3) for historical items, 7.2 (SD = 2.1) for the clinical items, and 5.6 (SD = 2.3) for the risk management items. Compared with non-violent subjects (M = 24.7, SD = 5.2), patients with aggressive behavior during the follow-up showed higher mean total scores on the HCR-20 (M = 29.8, SD = 4.8), t (72) = - 4.4, p < .001.

Point biserial (r_{pb}) correlations were reported for each of the instruments and violence (collapsed across physical violence towards others and physical violence towards objects) across 4-month intervals during the one-year follow-up period. The HCR-20 total score was found to correlate more significantly with violent behavior than the PCL: SV. HCR-20 total scores produced correlations with violence ranging between .35 and .45, with higher indices for the clinical subscale across all periods (r_{pb} between .49 and .54). The PCL: SV scores produced similar significant relationships with violence only in the first 4 months (r_{pb} = 0.36), and were non-significant after this period (r_{pb} was .18 and .21, respectively). After controlling for the PCL: SV the correlation between the HCR-20 and violence (r_{pb} = 0.45, p < .001) in the first four months dropped but remained significant (Partial r_{pb} = 0.32, p < .05). None of the PCL:SV correlations were significant after controlling for the HCR-20 suggesting that there was a unique relationship between the HCR-20 and violence that was independent of the variance in the HCR 20 attributable to the PCL:SV but not vice versa.

For violence (collapsed across physical violence towards others and physical violence towards objects), the AUCs of the HCR-20 total score (without item 7) were .75, .69 and .77 (all significant) for months 1 – 4, months 5 – 8 and months 9 – 12, respectively. The AUCs of the PCL: SV total score were .70, .63, and .61 (only the AUC value for the first four month period was significant) for months 1 – 4, months 5 – 8 and months 9 – 12, respectively. The HCR-20 final risk judgment (AUCs were .78, .78 and .77, respectively) and clinical subscale (AUCs were .77, .81 and .76, respectively) showed the highest predictive validity over the entire follow-up period.

The authors conducted a logistic regression to examine key predictors in relation to violence outcomes. The analyses used violent behavior for each period of follow-up as the dependent measure, and total scores and subscales of the HCR-20 and PCL: SV as predictors. When total HCR-20 and PCL: SV scores were included, only the HCR-20 contributed significantly to the prediction of physical violence during all periods. When HCR-20 subscales and PCL: SV factors were entered in a new set of analyses only clinical items remained in the model for each time period. The same analyses were repeated using physical aggression...
only as the outcome. The results were the same as above, the only exception being for 9-12 months risk management factors remained in the model but clinical factors did not.

The authors also conducted a logistic regression to examine whether structured final clinical judgments could add incremental validity to the model in addition to HCR numerical scores. These results revealed that for shorter time periods (i.e. the first eight months) structured final clinical judgments added incremental validity to the HCR-20 numerical scores used actuarially to predict violence in the hospital, but that after this period the numerical scores were robust single predictors of inpatient violence. These results were similar when the outcome was restricted to physical violence against persons.

This study demonstrated that the Spanish adaptation of the HCR-20 was a useful measure for predicting the likelihood of inpatient violence in a civil psychiatric environment in both the short term and over a one-year period. It should be noted that damage to property, if it is not fear-inducing behavior, does not match the definition of violence in the HCR-20 manual.

### Project Description

**Douglas, K. S., Isomura, T., & Koo, A. A Prospective, Repeated-Measures Study of Dynamic Risk Factors, Treatability, and Community Outcome among Civil Psychiatric Patients.**

This prospective study examines dynamic risk factors for violence, victimization, and self-harm in civil psychiatric patients admitted to an acute stay ward in a large hospital in Western Canada. Informed consent is obtained from participants through a detailed description of study procedures and a brief comprehension test. Baseline data is collected in the hospital via self-report, followed by an interview. Patients are re-contacted in the community for follow-up interviews at four-week intervals over a six-month period from the date of discharge. Follow-up information is also gathered from provincial correctional records. Violence, victimization, and self-harm are measured using the MacArthur Community Violence Instrument.

### Scholarly Works


### Summary

This study examined differences in HCR-20 scores, risk ratings, and predictive validity across several subgroups. The study sample consisted of 139 civil psychiatric patients admitted to an acute stay ward in Canada. Average age of the sample was 34 years. Most of the sample was male (53.2%) and Caucasian (78.4%). Primary Axis I diagnoses were mood disorder (68.1%), psychotic disorder (46.4%) and comorbid substance use disorder (34.1%). Subgroups were created by grouping participants by gender (male, female), diagnostic category (psychopathy, psychotic disorder, substance use disorder, antisocial personality disorder, mood disorder), history of violence (present, absent), and formal criminal contact (present, absent).

The authors found that patients with psychopathy scored significantly higher on HCR-20 total and subscales than patients without psychopathy. Differences were also observed at an item level, compared to patients without psychopathy patients with psychopathy scored higher on all but four items of the HCR-20. Compared to patients without a psychotic disorder, patients with a psychotic disorder scored higher on the C scale and five items of the HCR-20.

Patients with a substance use/abuse disorder scored higher on the total and H scales and 7 items of the HCR-20 compared to patients without a substance use/abuse disorder. Those with a history of formal crime contacts scored higher on H, C, and R scales and 9 items of the HCR-20 compared to those without a history of formal crime contacts. Patients with APSD scored higher on H and R scales, and 10 items of the HCR-20 then those without an APSD diagnosis. Last, patients with a mood disorder scored higher on 2 items of the HCR-20 and lower on 2 items relative to patients without a mood disorder. Differences in SPJ ratings across each of the subgroups were reported. SPJ ratings significantly varied as a function of gender, psychopathy, history of violence, history of formal crime contacts, APSD, and substance use/abuse.

The authors conducted logistic regression analyses to determine whether any of the sub-groupings moderated predictive validity of the HCR-20. The authors found that history of formal crime contacts moderated HCR-20 total, but not subscale, scores. None of the other subgroups moderated the predictive validity of the H, C, R or total scores. AUC values of the HCR-20 were reported for each of the subgroups and suggest that the HCR-20 predicted violence equally across diverse subgroups. Implications for risk management were discussed.

Summary

This study examined whether a history of homelessness improved the relationship between formal crime contact and the HCR-20, PCL: SV, Antisocial Personality Disorder module of the Structured Clinical Interview for the DSM Disorders (SCID), and the Criminal Sentiments Scale (CSS). Study participants were civil psychiatric inpatients (n = 117, 53.0% male) admitted to the acute stay psychiatric ward of a large hospital in western Canada. File information was obtained and participants completed an in-depth interview and self-report measures to determine if the participants had a history of any period of homelessness (NFA) and a history of formal crime contacts (arrests, charges, and convictions). Of the participants, 32.7% had a history of ever being homeless and 45.4% of the participants had a history of any arrests, charges, or convictions, with 38.9% having a history of non-violent contacts and 17.6% having a history of violent crime contacts.

The authors found that a history of NFA was significantly related to any formal contacts ($\chi^2 = 6.84, p < .01$), both non-violent and violent. A history of being NFA increased the odds of formal crime contacts by 2.99 times for any contacts, 3.24 times for non-violent contacts, and 4.85 times for violent contacts. With respect to non-violent contacts, NFA added incrementally to the relationships between the PCL: SV ($\Delta \chi^2 = 4.42, p < .01$), the CSS ($\Delta \chi^2 = 7.66, p < .01$), and the SCID module ($\Delta \chi^2 = 5.29, p < .01$) and non-violent formal crime contact, but not to the relationship between the HCR-20 and non-violent formal crime contact. With respect to violent contacts, NFA did not add incrementally to the relationship between violent formal crime contact and the PCL: SV, the HCR-20, and the SCID module, but it did add to the relationship between violent formal crime contacts and the CSS ($\Delta \chi^2 = 4.82, p < .01$).

Exploratory analyses revealed that homelessness performed as well as other validated risk factors. When NFA was entered in a backward stepwise logistic regression with the PCL: SV items, NFA remained in the final model along with item 8 (poor behavior controls). When entered in a backward stepwise logistic regression with all of the HCR-20 items, NFA remained in the final model along with items H2 (young age at first violent incident), H10 (prior supervision failure), and R4 (noncompliance with remediation attempts). The authors concluded that homelessness should be considered when assessing risk for violent and general offending.


Summary

The present research examined differences in symptomatology and risk profiles, using participants recruited from an acute stay psychiatric ward in a general hospital (short-stay) and various wards from a large tertiary psychiatric hospital (long-stay). The study sample consisted of 124 short-stay patients with an average stay of 24 days and 90 long-stay patients with an average stay of 2404 days. The short-stay inpatients (52.4% male) had a mean age of 34.36 ($SD = 10.43$). The majority were Caucasian (80.6%) or Asian (11.3%). The long-stay inpatients (65.6% male) had a mean age of 47.91 ($SD = 11.48$). The majority were Caucasian (78.7%) or Asian (9.0%). Participants completed a comprehensive interview and collateral file information was also obtained. Specific variables that were coded included official diagnoses, admission status, and GAF scores. The PCL: SV, HCR-20, START and BPRS were also scored.

Numerous differences were found between the short-stay and long-stay inpatients. Long-stay patients were more likely to have been in the hospital involuntarily ($\chi^2 = 22.87, p < .001$), were more likely to have a psychotic disorder ($\chi^2 = 53.65, p < .001$), and less likely to have a major mood disorder ($\chi^2 = 76.51, p < .001$) or an anxiety disorder ($\chi^2 = 7.30, p < .001$). No difference was found in the rates of substance related disorders. Additionally, long-stay patients evidenced lower mean GAF scores on admission (37.95 vs. 43.39) and higher mean BPRS scores (30.41 vs. 44.08). As a measure of onset of illness, the long-stay patients on average were younger when first hospitalized for psychiatric reasons (23.18 vs. 27.19). As well, long-stay patients were also more likely to have three or more prior psychiatric hospitalizations compared to the short-stay patients.

On average the long-stay patients scored higher on both the PCL: SV (10.36 vs. 6.97) and the HCR-20 (23.61 vs. 17.72). With respect to subscale scores, the long-stay patients received higher H scale scores (11.91 vs. 9.25), higher C scale scores (6.11 vs. 4.32), and higher R scale scores (5.51 vs. 4.24) than the short-stay patients. In addition, on the START the long-stay patients received higher vulnerability...
ratings (19.35 vs. 17.20) and lower strength ratings (13.48 vs. 20.53). Standard deviations were not reported.

With regards to the HCR-20 final risk judgments, the long-stay patients received higher risk ratings for future violence ($\chi^2 = 26.10, p < .001$). With regards to the START final risk judgments, differences were found on several ratings. Long-stay patients were judged to pose greater risk of going AWOL ($\chi^2 = 22.64, p < .001$) and engaging in self neglect ($\chi^2 = 24.12, p < .001$). In contrast, long-stay patients received lower ratings with regards to their likelihood of committing self harm ($\chi^2 = 10.13, p < .01$), attempting suicide ($\chi^2 = 37.99, p < .001$), and engaging in substance use ($\chi^2 = 15.47, p < .001$). Finally, no difference was found between the ratings of long and short-stay patients with regards to violence and victimization. The authors concluded that risk assessment tools were able to differentiate between long-stay and short-stay patients and aid in management decisions.

**Project and Scholarly Work**


**Summary**

This prospective cohort study investigated the validity of historical, dispositional, and clinical factors for predicting community violence in an acute mental health sample ($n = 114$) in the UK up to 20 weeks post-discharge. Baseline assessments were completed by researchers in the hospital and were based on an interview with the participant and a review of case records. Risk factors were measured using the following scales: HCR-20, PCL: SV, VRAG, Violence Risk Scale (VRS), Novaco Anger Scale (NAS), Barratt Impulsiveness Scale (BIS), Positive and Negative Syndrome Scale (PANSS), Brief Michigan Alcohol Screen Test (MAST) and Drug Abuse Screening Test (DAST). Violent behavior was measured with the MacArthur Community Violence Instrument using information collected through participant self-report, case records, and collateral information. In this study, violence was defined as any acts that included battery, sexual assaults, assaultive acts or threats made with a weapon.

The study sample had a mean age of 40.5 years. A majority of the sample was male (62.3%), Caucasian (91.2%) and had a primary diagnosis of schizophreniform disorder or mania-bipolar disorder (55.3%). Only 4.4% of the sample had a personality disorder primary diagnosis and 28.9% of the sample had a history of serious substance misuse. The mean length of inpatient stay was 77.8 days ($SD = 98.34$) and 82.5% of participants were discharged within 10 weeks of the baseline assessment. In the 20-week period post-discharge, 56 violent acts were recorded committed by 25.4% of the sample. There were a mean number of 1.9 violent acts per participant.

Inter-rater reliability was conducted on a subset of 20 cases for the historical items of the HCR-20 and the PCL-SV. ICC values were satisfactory for the historical items of the HCR-20 (0.97) and PCL: SV total (0.97). The inter-rater reliability between three raters based on 7 cases was also calculated for the VRAG, HCR-C, HCR-R, and VRS. ICC values were 0.99 for the VRAG, 0.85 and 0.83 for the clinical and risk management items of the HCR-20, and 0.96 for the VRS.

Mean scores on all the risk measures were significantly different between the violent and non-violent group. With regards to HCR-20 total scores, violent participants had significantly higher scores on the HCR-20 ($M = 15.24$, $SD = 8.08$) compared to non-violent participants ($M = 10.37$, $SD = 6.26$), $t (112) = -3.12, p < .01$. Violent participants ($M = 7.97$, $SD = 4.50$) also had higher scores on the historical subscale compared to non-violent participants ($M = 5.58$, $SD = 3.37$), $t (112) = -3.01, p < .01$. The same pattern of findings was also obtained for total scores on the VRAG, VRS, and PCL: SV, with violent participants scoring significantly higher on measures of risk.

ROC analyses revealed that all the risk scales significantly predicted post-discharge violence; however the HCR-20 total score had the largest (though comparably sized) AUC value of all the measures ($AUC = .67, p < .01$) and the HCR-20 Historical sub-scale was the only measure of historical factors found to significantly predict post-discharge violence ($AUC = .66, p < .05$). Using a median split at 10, the odds ratio of the HCR-20 for any violence was 3.02. The odds ratio was not reported for VRAG total scores ($AUC = .65, p < .05$) or VRS total scores ($AUC = .66, p < .05$). Other measures that were predictive of post-discharge violence were the NAS ($AUC = .68, p < .01$), BIS ($AUC = .66, p < .05$), PCL: SV Factor 2 ($AUC = .64, p < .05$), and PANSS Aggressive sub-scale ($AUC = .65, p < .05$). When controlling for age and gender in a logistic regression analysis, in each case the abovementioned significant scales and factors remained predictive of post-discharge violence.

The authors examined the association between the frequency of violence up to 20 weeks post-discharge compared with the independent variables. Frequency of violence significantly correlated with HCR-20 total score ($r = 0.44, p < .05$), but not VRAG ($r = 0.32, p > .05$) or the VRS risk measures ($r = 0.36, p > .05$). Of the disposition
factors, only PCL: SV total ($r = 0.39$, $p < .05$), interpersonal ($r = 0.37$, $p < .05$), and social deviance sub-scales ($r = 0.37$, $p < .05$) were significantly correlated with frequency of violence. The PANASS total ($r = 0.51$, $p < .01$), positive ($r = 0.51$, $p < .01$), and aggressive subscales ($r = 0.61$, $p < .001$) were also significantly correlated with frequency of violence. There was also a significant correlation between the clinical items of the HCR-20 and frequency violence ($r = 0.51$, $p < .05$), but neither the MAST ($r = 0.01$, $p > .05$) nor the DAST ($r = 0.07$, $p > .05$) significantly correlated with the frequency of violence post-discharge.

The authors concluded that risk scales found to be predictive of community violence in forensic samples were also predictive of post-discharge violence in acute mental health patients in England. In particular, the HCR-20 total was a significant predictor of risk, supporting the use of dynamic and static risk factors in violence risk assessment with this sample. The authors note that despite significant correlations between the HCR-20 and frequency of post-discharge violence, a high false positive rate (65%) was evident at the median split. Thus the predictive power of the HCR-20 was only moderate when compared with larger effects found in similar studies.

Project and Scholarly Work


Summary

From the authors’ perspective, research has demonstrated that the HCR-20 has moderate to strong predictive accuracy in men, however the utility of the HCR-20 has not been conclusively established in women. This retrospective study utilized data from the MacArthur Study of Mental Disorder and Violence to examine whether an abbreviated version of the HCR-20 was a valid indicator of violence risk in women. The five-item risk management scale and the historical item ‘prior supervision failure’ (H10) could not be scored due to missing data, and thus were omitted from subsequent analyses. As such, only 14 of the 20 items of the HCR-20 were scored. The dependent variable in this study was violence (coded as present or absent) from during the 20-week follow-up period. Violence, as defined by the MacArthur study (i.e. acts of battery that resulted in physical injury; sexual assaults; and assaultive acts that involved the use of a weapon; or threats made with a weapon), was coded as present if there was any physical violence to others by the discharged patient in the community.

Of the original 1,136 participants enrolled in the MacArthur study, 185 participants did not complete at least one of the five follow-up interviews and were excluded from the sample. An additional 124 patients were also excluded due to missing data on the PCL: SV and BIS-11, instruments which were used to code two items of the HCR-20. The final study sample consisted of 827 patients (477 males and 350 females) who had at least one follow-up. The average age of the participants was 29.8 ($SD = 6.2$). Of the sample, 68.7% were Caucasian, 29.1% were African America and 2.2% were Hispanic. Primary diagnoses were psychotic disorder (20.4%), unipolar mood disorder (42.8%) and bipolar disorder, mania or cyclothymia (12.1%). Men compared to women, were more likely to have a psychotic disorder and/or substance abuse diagnosis, significantly higher scores on the PCL: SV, more prior arrests, and have committed a higher number of violent acts. At least one incident of violence was reported for 155 individuals (18.7%) during the 20 week follow up period, committed by 22.2% of men compared to 14.0% of women.

To examine whether men and women would score differently on individual items of the HCR-20 the authors conducted an ANCOVA, controlling for race and age. Results indicated there were significant gender differences for eight of the 14 items analyzed. Men scored higher than women on previous violence (H1), substance use problems (H5), psychopathy (H7), and negative attitudes (C2). Women scored higher than men on relationship instability (H3), employment problems (H4), major mental illness (H6), and early maladjustment (H8). Despite significant differences at the item level, there was no significant main effect for gender (i.e., difference in average scores) for the HC ($M_{Men} = 14.09, SD_{Men} = 3.85, M_{Women} = 13.93, SD_{Women} = 3.54$), H ($M_{Men} = 9.68, SD_{Men} = 2.79, M_{Women} = 9.79 SD_{Women} = 2.59$), or C subscale scores ($M_{Men} = 4.42, SD_{Men} = 2.05, M_{Women} = 4.15, SD_{Women} = 2.04$) of the HCR-20.

For the sample as a whole, AUC values were moderate to poor (.66, .68, .54 for the HC total, H and C scales, respectively, all $p < .05$) There was no significant difference between HC and H AUC, however the combined HC total score was significantly more accurate in predicting violence than the C scale total score ($p < .001$). Likewise, the AUC was significantly higher for the H scale than the C scale ($p < .001$). When the H, C and HC scales were analyzed separately by gender, the HC total score yielded an AUC of 0.68 ($p < .001$) for men and 0.60 ($p < .05$) for women. The difference in AUC values was not significant. The AUC for the H scale was significantly greater for men ($AUC= 0.72, p <0.001$) than for women ($AUC=0.60, p <.05$). There was no difference between men and women for the C scale ($AUC = 0.54$ vs. $AUC=0.52$, respectively, both
ps > .05). AUC estimates were also reported separately for gender for each of the individual HC items. There were no significant differences in AUC estimates between men and women on any of the items.

Results suggested that the HCR-20 is slightly, but not significantly, better for evaluating future risk for violence in men than in women, although the magnitude of the gender differences was small and was largely limited to historical factors. In addition, HCR-20 ratings were made in an atypical manner, using proxies derived from a pre-existing data set that was not collected in order to inform ratings on the HCR-20. The authors concluded that the results do not indicate that the HCR-20 needs to be tailored for use in women or that it should not be used in women, but they do highlight that the HCR-20 should be used cautiously and with awareness of its potential limitations in women.

See Also


Project and Scholarly Work


Summary

The authors aimed to construct a brief checklist to use in civil psychiatric settings. The sample comprised all patients (N = 509) residing at a short-term inpatient unit in Oslo, Norway during a one-year period. The final sample consisted of the 110 patients for whom complete data were available. Participants were 55 women and 55 men whose mean age was 38.3 years (SD = 12.9, range: 19-77).

The authors constructed a 33-item “preliminary scheme” (PS) measure that consisted of all the HCR-20 items except H7 Psychopathy, 6 items from the BrØset Violence Checklist (BVC; Almvik, Woods, & Rasmussen, 2000), and 8 additional items based on the authors’ clinical experience and their review of the literature. PS items are scored using the same 0, 1, 2 criteria as the HCR-20. Physicians or psychologists responsible for each participant’s treatment completed PS ratings at discharge. Raters were trained in use of the measure. Protocols were excluded from analyses if there were missing data on more than six items. To assess interrater reliability, eight of the raters (the total number of raters was not specified) made independent ratings of the same 15 abridged real case stories. An ICC of 0.86 for the whole instrument was obtained.

Data on patients’ violence in the community was collected every three months over a one-year period. Information about violent outcome was based on patients’ self report during after-care consultations and “spontaneous information from family or friends.” Violence was defined as being verbally and/or physically violent towards others. Physical violence referred to any physical attack on a person. Non-physical violence was operationalized as threats to harm a person, verbal attacks and attacks on objects that could induce fear in a person nearby. However, all analyses were based on the aggregate “any violence,” which included violence of either type.

The mean total score of the 33-item PS was 15.9 (SD = 8.2; range: 4-42). Approximately one-quarter of participants (n = 29; 26%) engaged in at least one violent act during follow-up (M = 2.2, SD = 1.6; range: 1-7). There were 12 violent women and 17 violent men. Of the 29 violent patients, 13 (7 women, 6 men) had been physically violent; 14 (4 women, 10 men) had exhibited only verbally threatening behaviour; in two cases, the nature of the violent act was not specified.

Odds ratios (OR) for any violence for the 33-item PS ranged from 0.7 (HCR-20 R3 Lack of Personal Support) to 12.8 (“Present substance use”). The largest OR among the HCR-20 items was for H1 Previous Violence (OR = 7.0). The other items for which statistically significant ORs were obtained were: HCR-20 H2 Young Age at First Violent Incident (OR = 3.8); HCR-20 H5 Substance Use Problems (OR = 2.9); HCR-20 H10 Prior Supervision Failure (OR = 2.8); HCR-20 C1 Lack of Insight (OR = 2.7); BVC item Verbal Threats (OR = 4.8); BVC item Physical Threats (OR = 5.0); “Suspicousness” (OR = 2.7); “Lack of Empathy” (OR = 3.3); HCR-20 R1 Plans Lack Feasibility (OR = 2.4); and HCR-20 R5 Stress (OR = 3.6).

AUC values associated with engaging in any violence were 0.71 (p < 0.01) using the 33-item PS and 0.73 (p < 0.01) using the 19 HCR-20 items. AUC values higher than the 0.71 associated with the whole PS were obtained when various combinations of items with significant ORs were used. More specifically, combinations of 4, 6, and 8 items yielded AUCs of 0.77, 0.77, and 0.76, respectively (all p < 0.01). The authors argued that their data support the possibility of developing a brief screening instrument specifically for use in acute psychiatric units.
Summary

As the HCR-20 is used on various psychiatric populations in practical settings, this study investigated the accuracy of the HCR-20 in a sample of veterans suffering from posttraumatic stress disorder (PTSD). This population is known to exhibit hostility, negative affect and violent behaviour, as well as increased risk for violence when compared to control samples.

Using a prospective design, 104 male veterans diagnosed with PTSD were assessed using the HCR-20. Participants had a mean age of 35 years (SD = 10.2), most of whom were married, employed, had children, and previously competed secondary school. Violent outcomes were assessed bimonthly for a one year period. Of the participants, 67% perpetrated at least one violent act in the follow-up period, while 56% perpetrated at least one act of physical violence.

Interrater reliability was assessed on a subset of 52 participants. ICCs are reported for the total score (ICC = .95) and for each subscale separately (H = .96, C = .75, and R = .88). The ICC for the final risk judgments was .88.

Predictive validity was examined in several manners. The authors report AUCs for any violence, non-physical violence and physical violence, each separated by subscale and total score. For any violence, AUCs were .79, .85, .83, .70, and .71, for the final risk judgments, total score, H scale, C scale, and R scale, respectively. For non-physical violence, the AUCs were .73, .82, .81, .70, and .69, again for the final risk judgments, total score, H scale, C scale, and R scale, respectively. Finally, for physical violence, the AUCs were .76, .86, .86, .73, and .69 again for the final risk judgments, total score, H scale, C scale, and R scale, respectively.

Additionally, for each of the violence indexes a logistic regression was conducted to determine which items were independently significant predictors of violence. With regards to the perpetration of any violence, three variables were found to be predictive: R3 (lack of personal support) with $e^b = 2.389$, H1 (previous violence) with $e^b = 3.8642$ and R5 (stress) with $e^b = 2.597$. With regards to the perpetration of non-physical violence, three variables were found to be predictive: R3 (lack of personal support) with $e^b = 2.676$, H1 (previous violence) with $e^b = 12.898$ and C5 (unresponsive to treatment) with $e^b = 2.182$. With regards to the perpetration of physical violence, three variables were found to be predictive: C1 (lack of insight) with $e^b = 3.851$, H1 (previous violence) with $e^b = 10.398$ and R5 (stress) with $e^b = 2.374$.

The authors concluded that the results support the use of this instrument with this population, as the results were comparable to those of other forensic psychiatric, civil psychiatric and correctional samples.

Project and Scholarly Work


This prospective study examined the predictive validity of the HCR-20 and PCL: SV for violence and factors related to the manifestation of violent behavior (re-hospitalization, aggressive behavior, suicide attempts) over a three-year period in a sample of civil psychiatric patients in Greece. Patients were included if the study if they were between the ages of 18 and 70 and exhibited some form of violent or aggressive behavior prior to their current hospitalization. The final sample consisted of 295 (159 male and 136 female) individuals. Mean age of the sample was 41.4 years. A majority of the sample was single (68.5%), unemployed (40.3%), and had been diagnosed with schizophrenia/psychotic disorder (46.1%). Other diagnoses in the sample were personality disorder (11.2%) and co-morbid substance (20.7%) and alcohol use (22.4%).

A week prior to discharge patients were assessed with the HCR-20, PCL: SV, and GAF using information collected from case file, interviews with patients, clinical staff, and collateral informants. During the follow-up period, outcome information was collected from patients and collateral informants at 6-month intervals. At the time of the initial assessment, mean scores on the HCR-20, PCL: SV, and GAF were 28.3 (SD = 4.4), 13.4 (SD = 4.7), and 48.4 (SD = 10.3), respectively. At the time of the second assessment, mean scores were 29.8 (SD = 3.0) and 14.4 (SD = 4.3) on the HCR-20 and PCL: SV, respectively. Total scores on the HCR-20 and PCL: SV were significantly positively correlated ($r = .61, p < .001$), in addition all factor and subscale scores were significantly positively correlated. Internal consistency, using Cronbach’s alpha, was .70, .72, .65, .71 for Total, H, C, and R scales, respectively. Cronbach’s alpha was not reported for the PCL: SV.
During the three year follow-up period, 44.4% of the sample were re-hospitalized. Although rates of violence were not reported, the authors note that the most common form of violence was aggression towards others (which occurred in 82.4% of participants who were violent). Rates of suicide (attempts, completion) were not reported. Both the HCR-20 and PCL-SV total and subscale scores were found to be significant predictors of hospital readmission, suicide attempts, and violent behavior. With respect to re-hospitalization, AUC values were .57, .57, .63, .59, .60, and .59 for PCL: SV Total, Factor, Factor 2, HCR-20 Total, H, C, and R scales, respectively. All scales except Factor 2 of the PCL: SV were significant. With respect to suicide attempts, AUC values were .53, .56, .54, .68, .57, .62, and .70 for PCL: SV Total, Factor, Factor 2, HCR-20 Total, H, C, and R scales, respectively. All scales except the HCR-20 C scale were significant. With respect to violence, AUC values were .66, .66, .61, .68, .65, .60, and .63 for PCL: SV Total, Factor, Factor 2, HCR-20 Total, H, C, and R scales, respectively. All scales were significant. Results of a series of logistic regressions indicated that increases on the HCR-20 Total scale significantly increased the probability of readmission to the psychiatric unit, probability of successful suicide, and aggressive behavior.

This study was the first in Greece to test the validity of the HCR-20. The authors concluded that the results provide strong evidence that the HCR-20 is a good predictor of violent behavior in psychiatric patients, and therefore should be used by clinicians in routine clinical practice in Greece.

**Project and Scholarly Works**


**Summary**

These authors used a pseudo-prospective design to evaluate the utility of three decision support tools for assessing acute risk of violence: the HCR-20, the PCL-SV, and the McNiel-Binder Violence Screening Checklist (VSC). 100 patients from a university-based, short-term psychiatric inpatient unit were used as participants. The design used a case-control method of sampling in which 50 individuals who had been physically assaultive were matched with 50 cases who had been nonviolent patients. For this study, the definition of violence was operationalized as physical attacks on persons. The median length of hospitalization was 9.5 days.

Inter-rater reliability as measured by ICC for the devices were: HCR-20 = .78, PCL-SV = .77, VSC = 1.0. The means from the study group were: HCR-20 total (SD = 6.6), HCR-20 H-scale 7.1 (SD = 3.5), HCR-20 C-scale 6.1 (SD = 2.3) and HCR-20 R-scale 4.8 (SD = 2.3), VSC 2.1 (SD = 1.3), PCL-SV total score 9.1 (SD = 5.1), PCL-SV Part 1 4.7 (SD = 3.0) and Part 2 4.5 (SD = 2.8).

Correlational analyses showed that the HCR-20 total score was correlated with the PCL-SV total score (r = .61; p < .01) and with the VSC (r = .26; p < .01). Each of the HCR-20 scales is also correlated with the PCL-SV total score (H-scale (r = .56; p < .01), C-scale (r = .4; p < .01) and R-scale (r = .47; p < .01)) and with the VSC (H-scale (r = .17; p < .01), C-scale (r = .34; p < .01) and R-scale (r = .15; p < .05)).

Logistic regression analyses showed that when violence was predicted based on the total scores from the PCL-SV, the HCR-20 and the VSC, that only the VSC made an independent contribution to the violence prediction. Further regression analyses showed that when violence was predicted based on the subscale scores from the PCL-SV, the subscales of the HCR-20 and the VSC, that the Clinical items from the HCR-20 and the VSC made independent contributions to violence prediction.

ROC analyses of the HCR-20 subscales showed AUC’s of .56 for the H Scale, .77 for the C Scale and .58 for the R Scale. For the PCL-SV, the AUC for Part 1 was .66 and for Part 2 was .55. Of these subscales, only the HCR-20 Clinical items and Part 1 of the PCL-SV differed significantly (p < .01) form the line of no information. Compared to research using the HCR-20 with long-term community follow-up, the HCR-20 had generally lower levels of sensitivity and specificity in this sample.

The discussion section details the need for risk assessment tools as well the need for tools that are more appropriate for short-term risk assessment as opposed to long-term risk assessment. However, the C scale of the HCR-20 was shown to be an important independent predictor of short-term inpatient physical violence.


**Summary**

This study compared the accuracy of risk assessments completed by experienced psychiatrists with those completed by psychiatric residents. The study used a retrospective case-control design. Medical records were
programs were reviewed for details of the risk assessments standards according to the authors. Patients’ files in both of a short check list that falls with their standard risk assessment process, which consisted of a one day refresher training after four weeks. A control group was provided with standard two day HCR training and a one day refresh training after four weeks. A control group was provided with standard two day HCR training and a one day refresh training after four weeks.

This study aimed to examine the effect of implementing the HCR-20 into an Assertive Outreach Community Mental Health Team. The PATHWAYS Project: Evaluating the transition of psychiatric services from hospital to community.

Project and Scholarly Work


Summary

This study aimed to examine the effect of implementing the HCR-20 into an Assertive Outreach Community Mental Health Team. One team of mental health professionals was provided with standard two day HCR-20 training and a one day refresher training after four weeks. A control group used their standard risk assessment process, which consisted of a short check list that falls below professional standards according to the authors. Patients’ files in both programs were reviewed for details of the risk assessments and violent or antisocial outcomes. Over the 12 month period, the incidence rate of violence in the trained group was 41% compared to only 20% in the control group. Similarly, the incidence rate of contact with police was 59% in the trained group and 37% in the control group. Therefore, there were more violent acts in the trained group. However, it was found that only 2 risk assessments using the HCR-20 were actually conducted in this period by the trained group. Instead, the trained group continued to use the standard check list. Both groups did however generate risk scenarios, or case conceptualizations, for the majority of cases. Nevertheless, these risk scenarios did not contribute to risk management. Overall, the authors concluded that training alone is inadequate to change the risk assessment process utilized in such programs, and that in order for a structured tool to be implemented into clinical practice a “cultural shift” in the mental health staff must occur.

Project Description

Nicholls, T. The PATHWAYS Project: Evaluating the transition of psychiatric services from hospital to community.

This study examines the practical, clinical, and social implications of transferring chronically ill patients from a large psychiatric hospital in Western Canada to community-based settings. Prior to transfer, a detailed evaluation of each patient’s clinical (e.g. physical health, psychiatric symptoms), behavioural (e.g. suicide, self-harm, aggression, activities of daily living), and psychosocial (e.g. consumer satisfaction, quality of life, stigma) status was conducted. After moving into a community care setting, each patient was re-assessed several times to determine what, if any, changes were found. In addition to patient interviews, information from patients’ family members and peers, as official record databases were also used in the study.

Scholarly Works


Summary
The Female Additional Manual (FAM) supplements the HCR-20 with variables theoretically relevant to variables specifically relevant to women’s violence, though the tool has yet to be widely researched. This 6-month prospective study contributes to cross-validation research by examining gender differences in HCR-20 performance (descriptively and predictively) in a chronically ill psychiatric sample transferring to the community following a hospital closure.

The study sample consisted of 106 (65 men and 41 women) patients. The average age of the sample was 47.1 (SD = 12.4). A majority of the sample was Caucasian (75.5%) and had a diagnosis of Schizophrenia (31.1%). Additionally, 33% had a dual substance abuse diagnosis and 33% had a dual personality disorder or traits. For each participant the HCR-20 and START-Outcome Scale (START-SOS) were coded. The START-SOS assesses incidents of physical, verbal, and sexual aggression. Responses are coded on a 1-4 severity scale. In this study, START-SOS ratings were collapsed across file review, staff interview, and patient interview.

The authors found that men scored significantly higher on young age at first violence (H2) and women scored significantly higher on stress (R5). There were no significant differences between males and females on the H (12.47 vs. 11.02), C (6.1 vs. 6.5), R (5.7 vs. 5.5) or Total (23.9 vs. 22.9) scales (standard deviations were not reported). Final risk judgments were also similar between men and women (30% vs. 27% were rated as low risk, 48% vs. 44% were rated as moderate risk, and 22% vs. 29% were rated as high risk). There were no gender differences in aggression type and severity over the follow-up period.

The authors examined gender differences in predictive validity of the HCR-20 for verbal, physical and overall aggression using Receiver Operating Characteristics (ROC). For men, AUC values of total and subscale scores ranged between .55 and 0.69 for verbal aggression, .38 and .60 for physical aggression, and .61 and .69 for overall aggression. For women, AUC values of total and subscale scores ranged between .43 and .66 for verbal aggression, .37 to .61 for physical aggression, and .43 to .66 for overall aggression. Differences in bivariate effect sizes hinted at better prediction for men, however there was no moderating effect of gender. Collapsed across gender, the results suggest relatively poor performance of the HCR-20 overall (AUCS ranged between .52 to .64 for verbal aggression, .47 and .58 for physical aggression, and .56 and .63 for overall aggression) however the authors suggest that high levels of intervention with study participants may have influenced findings.

Petersen, K.L., Douglas, K., & Nicholls, T. (2020, March). Dynamic nature of risk assessed by the HCR-20 in a long-term civil psychiatric population... project description...
Summary

This was a chart review study of all 279 involuntarily committed persons from a large psychiatric hospital in Western Canada who applied for Review Panel hearings in 1994. Data were collected concerning patients’ demographic characteristics, family and childhood history, mental health history, criminal history, and Review Panel hearing outcomes. The majority of patients had psychotic disorders, previous psychiatric hospitalizations, and were unemployed at admission. Over half of patients had previous arrests or convictions. Patients were tracked in the community after their release for an average of 2 years. Follow-up information was gathered from rehospitalizations to the releasing psychiatric hospital, hospitalization records from 16 general hospitals in the province, provincial correctional records, and Coroner’s records.

Scholarly Works


Summary

This study compared the predictive validity of the HCR-20 Risk Assessment Scheme (Webster, Douglas, Eaves, & Hart, 1997a; Webster, Eaves, Douglas, & Wintrup, 1995) and the Psychopathy Checklist: Screening Version (PCL:SV; Hart, Cox, & Hare, 1995). This research includes the 193 patients for whom complete measures were attainable (HCR-20; PCL:SV). Patients were followed into the community for an average of 626 days.

Violence was defined to include a demarcation between physical and non-physical aggression. Physical aggression refers to any attacks on persons. Non-physical aggression includes threats to harm a person, verbal attacks on persons, and “fear-inducing” behaviour such as attacks on objects. Violent crime was coded separately to allow for additional analyses, although typically it would also be coded as physical violence. The three types of violent outcome, then, were (1) any violence; (2) physical violence; (3) violent crime.

The AUCs produced by ROC ranged from .76 (for any and physical violence) to .80 (for violent crime). Odds ratios showed that persons scoring high on the HCR-20 (above the median) were 6 (for any and physical violence) to 13 (for violent crime) times more likely to be violent in the community than persons who scored under the median.

For the PCL:SV, AUCs ranged from .68 (for any violence) to .73 (for physical violence) to .79 (for violent crime). Effects for the PCL:SV were more variable than those for the HCR-20. The odds of violence for those above the median score of the PCL:SV also increased substantially (from approximately 4 to 13 times).

Hierarchical regression analyses revealed that the HCR-20 added to the predictive validity of the PCL:SV, but the converse was not true. Multiple regression analyses of the subscales of the HCR-20 and PCL:SV indicated that only HCR-20 scales predicted rate of violence. The H scale and R scale of the HCR-20 produced the largest effect sizes of all subscales with violence. Implications for research on risk assessment, as well as the clinical assessment and management of violence, are discussed.


Summary

This presentation focused on risk for inpatient violence specifically. Inpatient violence was defined in a similar manner as community violence. A distinction was made between physical violence (which required physical contact with victims) and non-physical violence (which included threats of violence and fear-inducing behaviour). Approximately half of patients displayed physical aggression while hospitalized.

AUCs for the H and C Scales composite for inpatient violence ranged from .57 to .65. Odds ratios for inpatient violence averaged approximately 2.0, and for repetitive inpatient violence, 3.0. These values are smaller than those for community violence, and indicate a moderately sized relationship between the HCR-20 and repetitive inpatient violence. The AUCs for the PCL:SV for inpatient violence were similar to those for the HCR-20, ranging from .60 to .64. Odds ratios were comparable to those of the HCR-20, averaging approximately 1.75 for inpatient violence, and 3.0 for repetitive violence.


Summary

The focus of this research was to compare the performance of the HCR-20, PCL:SV and a violence screening measure for both civil psychiatric inpatient and community violence. Concerning violent and criminal behaviour, a greater proportion of men had histories of crime (including violent crime). On the violence outcome measures, there were no differences in the incidence of inpatient violence across genders. A greater percentage of men compared to women displayed community violence.

Males had higher mean scores on the H Scale ($M = 10.8; SD = 3.3$), C Scale coded upon admission ($7.4; SD = 1.5$), and HCR-20 Total Scores ($M = 20.4; SD = 5.6$) compared to women (H Scale $M = 8.2; SD = 3.2$; C Scale at admission $M = 6.9; SD = 1.7$); HCR-20 Total Score $M = 16.8; SD = 5.4$). Males also had higher scores on the PCL:SV.

ROC AUC values for inpatient violence showed that the HC composite, the PCL:SV, and McNiel and Binder’s (1994) Screening Measure did not predict violence for males. However, for females, moderate to large effects were observed for the HC composite (AUCs = .62-.74) and the PCL:SV (AUCs = .63 = .74). McNiel and Binder’s screening measure was weakly related to violence, predicting only verbal aggression.

For post-release community violence, a different picture emerged. For male patients, HCR-20 AUCs ranged from .72 (any violence) to .73 (physical violence) to .75 (violence resulting in criminal sanctions). PCL:SV AUCs ranged from .63 (any violence) to .70 (violence resulting in criminal sanctions) to .71 (physical violence). For females, HCR-20 AUCs ranged from .66 (physical violence) to .77 (any violence) to .80 (violence resulting in criminal sanctions). PCL:SV AUCs ranged from .51 (physical violence) to .67 (any violence) to .89 (violence resulting in criminal sanctions).

This study is important because it focuses on gender. Perhaps surprisingly, it found that the HCR-20 performed better for the prediction of inpatient violence by women than by men. Prediction of community violence was comparable between genders. Statistical comparisons were not made between genders or measures, and as such the differential predictive validity was not addressed directly. Further, analyses were not carried out for HCR-20 and PCL:SV subscales.

See Also


Project and Scholarly Work


Summary
The objective of this prospective study was to analyze clinical factors involved in violence by using the HCR-20, the Global Impression of Severity (CGI-S), and the Clinical Global Impression of Improvement (CGI) over a follow-up period of 6 to 12 months in a sample of males with schizophrenia. The study sample was comprised of 59 patients admitted to an acute psychiatric ward in Romania. The identification of violence during follow-up was completed through a non-structured psychiatric interview and the collection of collateral data from the patient’s family. Violence’s severity was categorized as mild (e.g., threats), moderate (e.g., beating) or severe (e.g., sexual offense, rape, homicide). Of the sample, 30.5% committed mild violence and 15.3% committed medium violence. No severe violence was observed in the sample.

The HCR-20, PCL-R, CGI-S and CGI were coded the first week after admission and again a minimum of 6 months after the initial assessment. The CGI-S and the CGI analyze the severity of disease and the evolution under treatment, respectively. Scores on the CGI-S and CGI indicate that a majority (80%) of the sample had a partial response to treatment over the 6-month interval between assessments. Scores on the C subscale of the HCR-20 were also found to significantly decline between the two time points. At time point one, mean scores were 19.41 (SD = 5.59) 7.05 (SD = 3.03), 6.83 (SD = 1.66), and 5.58 (SD = 2.43) on the total, H, C and R scales, respectively. At time point two, mean scores were 18.31 (SD = 5.65), 7.56 (SD = 3.01), 5.68 (SD = 1.93), and 5.24 (SD = 2.38) on the total, H, C and R scales, respectively. There were no significant changes to the H, R or total scores of the HCR-20 over the 6-month interval.

Correlation analysis between select items of the HCR-20 (C1 – C5, C total score and R4), adherence to treatment, severity of violence, and the prediction of risk for violence indicated that adherence to treatment was significantly related to the prediction of future risk for violence (r = .41, p < .001). C total scores were significantly related to adherence to treatment (r = .35, p < .001) and prediction of the risk for violence (r = 0.49, p < .001) but were uncorrelated with the severity of the violent act. Items C1, C2, C4, C5 and R4 were also found to be significantly related to the prediction of risk for violence, but only C1, C2, C5 and R4 were significantly related to adherence to treatment.

Project and Scholarly Work


Summary

This study examined the reliability and validity of the HCR-20 for assessing violence risk in a sample of male psychiatric patients in China. Participants were sampled from male patients who were admitted to hospital between September 2008 and December 2008. Thirty participants were selected on the basis of past aggression, defined as any intentional action against other people which resulted in at least minor injuries (aggressive group). Of these participants, 27 had committed homicide and 3 had caused severe injuries. Another thirty participants were selected from the same period, but did not have any past incidents of aggressive behavior (non-aggressive group). All male participants were diagnosed with schizophrenia. Aggressive and non-aggressive groups did not significantly differ with respect to age, marital status, educational level, or treatment.

Using information collected from self-report, hospital files and interviews with patients and their close relatives, researchers completed the HCR-20, BIS – 11, and BPRS. Internal consistency, using Cronbach’s alpha, was .82, .91, .78 and .87 for Total, H, C, and R scales on the HCR-20, respectively. Cronbach’s alpha was not reported for the other measures. Test-retest reliability of the HCR-20 was determined by re-coding the HCR-20 two to three weeks after the initial assessment. Test–retest reliability of HCR-20 Total score was .90. The Modified Overt Aggression Scale (MOAS) was used to determine patient aggression over the past month. Four types of aggression were considered: verbal aggression, aggression towards property, autoaggression (i.e., self-injurious behavior) and physical aggression. In the past month, 45.5% of the aggressive group and 15.5% of the non-aggressive group committed acts of aggression while in the hospital. With respect to the aggressive group, 36.5% were verbally aggressive, 41.2% were physically aggressive, 37% were aggressive towards property and 45% were autoaggressive. With respect to the non-aggressive group, 24.6% were verbally aggressive, 19.8% were physically aggressive, 24% were aggressive towards property, and 16.0% were autoaggressive. All differences between the two groups were significant (ps <.01).

Participants in the aggressive group (M = 80.37, SD = 8.00) scored significantly higher on the BIS-11 compared to non-aggressive participants (M = 62.83, SD = 10.95), t (59) = 7.84, p < .01. There were no significant differences between participants in the aggressive and non-aggressive groups on BPRS total and subscale scores. With respect to the aggressive group, mean scores were 57.50 (SD = 10.64),
Project and Scholarly Work


Summary

This study assessed the feasibility and clinical utility of the HCR-20 in general adult psychiatric admissions. During the 5-month study period, 144 patients were admitted to one of two general adult wards and 135 (75 males and 58 females) participated in the study. The HCR-20 (without the Psychopathy item due to time concerns) was completed by staff based on the medical and nursing notes from admission and an interview with the participant if the information was incomplete.

There was no difference between men and women with regards to their ages (males, \(M = 37.8\) years, females \(M = 38.1\) years), age at first symptoms and previous number of admissions. Male patients were more likely to have a primary diagnosis of schizophrenia and acute psychosis (males, 45.9%; females, 18%), and were more likely to have a comorbid diagnosis of drug and alcohol use (males, 43.9%; females, 21.6%), whereas female patients were more likely to have a primary diagnosis of affective disorder (males, 31.7%; females, 57.4%).

The average HCR-20 total score for all patients was 18.0 \((SD = 7.3)\). There was a significant difference between males and females for total scores \((M = 19.2, SD = 6.4; M = 16.0, SD = 8.2\) respectively). Participants in the high-risk group \((n = 28; HCR-20 > 25)\) were just as likely to be female as male than the low risk group \((n = 105)\). They were more likely to have a primary diagnosis of personality disorder, much less likely to have an affective diagnosis, and more likely to have a comorbid substance problem. The HCR-20 total score, the H subscale and the R subscale scores were highest in those patients with a diagnosis of personality disorder. The C subscale was similar across diagnoses. In terms of length of stay in the hospital, patients who were in the hospital for less than 10 days had a significantly high H subscale score and those who stayed longer had a significantly higher C subscale score.

Logistic regression revealed that the total HCR-20 score did not predict length of stay but a diagnosis of personality disorder predicted a short stay in hospital. This remained significant when the HCR-20 total score and being defined as high risk were included in the model. In terms of feasibility, it was possible to complete HCR-20s on 83.9% of admissions within 24-48 hours of admission.

Project Description

Webster, C. D., Hart, S. D., Eaves, D. Prospective study of the HCR-20 in a civil psychiatric setting.

Summary

This was a prospective study of 131 persons admitted consecutively to the Intensive Care Unit (ICU) of a large psychiatric hospital in Western Canada. There were 82 (63%) men and 49 (37%) women. The mean age at admission was 36 years \((SD = 12)\). The majority of patients were single \((n = 105; 80%)\). Only 10% \((n = 13)\) of the sample was employed at admission. The mean length of stay on the ICU was 21 days \((SD = 12)\). Patients had on average 6.1 \((SD = 6.4)\) previous psychiatric hospitalizations. Over half of the sample had schizophrenic or other psychotic
disorders as admission diagnoses ($n = 73$; 56%). Approximately one-fifth ($n = 28$) of the sample received diagnoses of personality disorder.

The HCR-20, PCL:SV, and BPRS were completed for each patient. Research assistants coded the H scale items, and attending psychiatrists coded the C and R scale factors. Violence was measured on the unit by use of the Overt Aggression Scale. Patients were also tracked in the community. Subsequent contacts with corrections, police, and hospitals were recorded from archival sources. A research assistant also contacted community “collaterals” (persons who knew the patients and could report on their community behaviour) at three and six months post-release.

Scholarly Works


Summary

In a subset of this sample comprising 50 patients, the 10 Historical variables of the HCR-20 and the 12 items from the PCL:SV were used to predict inpatient violence. Violence included acts of verbal aggression, self-directed aggression, and aggression toward others and objects (as measured by the Overt Aggression Scale). With respect to internal consistency of the HCR-20 H scale, Klassen reported a Cronbach’s alpha of .73. Correlations between the H variables and violence averaged .30 across several outcome measures, and controlling statistically for the effects of sex. Of the individual items, substance abuse and psychopathy were most strongly related to violence. The PCL:SV performed similarly to the H Scale, correlating at .26 with ward violence. Part 2 of the PCL:SV, which measures the behavioural aspects of psychopathy, was somewhat more strongly related to ward violence (.33) than were PCL:SV Total or H scores from the HCR-20.


Summary

In this study, 47% ($n = 62$) of patients displayed violence toward others while hospitalized. For inpatient violence, the H Scale, C Scale, and HC composite produced AUCs with violence that were greater than chance, ranging from .63 to .68 for any type of aggression. The largest AUC was for the HC composite. The PCL:SV AUC was .61. The HCR-20 H and C scales were related to ward violence with moderate strength in this sample Survival analyses showed that persons who scored high on the HC composite were twice as likely (62%) to be violent by day 10 post-admission compared to persons who scored low (35%).

For the community phase of the study, 112 patients had been released by the end of the study period, and data were complete for 101 of these patients. Half of the sample displayed violent behaviour in the community, most frequently verbal aggression to others. For the HCR-20 subscales, AUCs for any aggression to others ranged from .58 (C) to .73 (R). For physical violence, the AUCs averaged approximately .63. The AUC for the HCR-20 Total score was .67. For violent crime, however, the HCR-20 AUC was .75. For the PCL:SV, the AUC for any violence and physical violence was .65, and for violent crime, .70. All AUCs are significantly greater than chance prediction.


Summary

This prospective study was done to examine the predictive validity of the HCR-20, PCL:SV, VSC and BPRS with regard to institutional aggression. A specific focus of the current project was to examine if dynamic risk factors contribute uniquely to the risk assessment process.

Participants were 131 psychiatric inpatients that were consecutively admitted to a large psychiatric hospital. The participants were on average 36 years old ($SD = 12.24$) and mostly male (63%). The H scale was scored based on file information by researchers, whereas the C scale was scored by the psychiatrist after an interview and file review. The VSC and PCL:SV were rated based only on file information. The BPRS was rated by the psychiatrist. Inpatient aggression was assessed using the Overt Aggression Scale (OAS) based on the patients’ files, nurses’ notes, and interviews with staff. Of the patients, 53% perpetrated at least one act of aggression and 30% perpetrated at least one act of aggression against people.
Predictive validity was examined with AUCs and correlations. With regards to any aggression, the results were as follows: H scale (AUC = .63, $r = .21$), C scale (AUC = .65, $r = .27$), HC (AUC = .68, $r = .28$), VSC (AUC = .64, $r = .28$), PCL:SV (AUC = .62, $r = .21$), and BPRS (AUC = .69, $r = .31$). With regards to any physical aggression towards people, the results were as follows: H scale (AUC = .59, $r = .14$), C scale (AUC = .60, $r = .20$), HC (AUC = .63, $r = .19$), VSC (AUC = .62, $r = .21$), PCL:SV (AUC = .59, $r = .15$), and BPRS (AUC = .71, $r = .32$).

Incremental validity analyses were performed to determine if the dynamic risk factors added to the static risk factors. Hierarchical logistic regressions were used in which the PCL:SV, H scale, and VSC were entered in the first block. This model was not significantly predictive of any aggression over chance levels. Then the C scale and BPRS were entered in the second block. The addition of these dynamic factors resulted in a significant model, with only the BPRS being a significant predictor individually. These analyses were also done with physical violence as the dependent variable and the same results were seen.

The authors concluded that these instruments can be used to predict inpatient aggression, although the effect sizes seen were quite small. Moreover, the instruments were better at predicting any aggression, compared to physical aggression. Finally, the (dynamic) factors, as measured with the BPRS, added incremental validity over the historical or static factors.

**Project and Scholarly Work**


**Summary**

The current study examined how psychiatric residents perceive their ability to conduct violence risk assessments and the risk factors they considered relevant in an emergency department setting. The authors surveyed 55 psychiatric residents at a Canadian university about their experience and education in assessing suicide and violence risk. The residents and a comparison group of 11 staff psychiatrists at a teaching hospital affiliated with the university then participated in a mock interview with one of the study authors. The subjects were directed to ask for all risk factors that would be relevant in determining the risk of a hypothetical patient with homicidal ideation. The risk factors they requested were compared with the risk factors found in the HCR-20.

On average, residents asked for 8.5 risk factors on the HCR-20 compared with 14.7 by the staff. Staff psychiatrists and residents were compared on the number of HCR-20 risk factors they asked for in the interview using between-groups ANOVA. Overall, the groups significantly differed in the number of risk factors they asked about with staff psychiatrists asking about more risk factors than residents. Further, junior residents asked about significantly fewer risk factors than senior residents.

The authors found that number of HCR-20 items requested significantly correlated with years of training, amount of formal and informal education, the number of patients for whom participants had discharged a duty to warn, and the number of suicidal and violent patients previously assessed. Confidence at assessing violence risk was not correlated with performance.
Summary

This study investigated the interrelationship and overlap between the Camberwell Assessment of Need–Forensic and the Health of the Nations Outcome Scales–Secure. Participants were 72 forensic psychiatric patients residing in a state forensic mental health facility in Australia. The study included 66 males and 6 females, aged between 20 and 62 years of age ($M = 37.78$, $SD = 8.98$) who had committed serious violent offenses (i.e., murder, attempted murder, serious assault). A majority (68%) of the sample had schizophrenia as their primary diagnosis.

The CANFOR-S was used to assess both patient and nurse views of total, met, and unmet needs across 25 domains of individuals experiencing mental illness who have offended. Each domain is coded on a three-point scale from 0 (need is present) to 2 (need is unmet). The HoNOS-S was also used to assess the needs of individuals experiencing mental illness who have offended. It comprises amended versions of the original 12 items HoNOS items and an additional seven-item security scale. Each item is rated on a five-point scale (0-4). A security scale item rated 1 or above indicates that a risk management intervention is needed and for items on the clinical scales, a rating of 2 or above indicates that a care or treatment intervention is needed. The HoNOS-S on the other hand was correlated with specific violence risk factors associated with mental illness rather than broader criminogenic needs. Thus, these two measures contributed different information to care and treatment planning.

Pearson’s product moment correlations were used to examine the relationship between each of the four measures. The HoNOS-S and CANFOR-S were found to have moderate to strong associations with the HCR-20. CANFOR-S nurse and patient ratings of total needs were significantly and positively correlated with the HoNOS-S clinical and security scales ($r_s 0.57$ and .79), as well as the HCR-20 clinical ($r_s 0.45$ and .68) and risk management scales ($r_s 0.69$ and 0.73). HoNOS-S clinical scales were also significantly, positively correlated with clinical ($r = 0.71$) and risk management scales of the HCR-20 ($r = 0.45$). HoNOS-S security scales were only found to be significantly, positively correlated with the clinical scale of the HCR-20 ($r = 0.67$). The HCR-20 was the only measure found to be significantly correlated with the LS-CMI ($r = 0.75$), correlations between the LS-CMI and the CANFOR or HoNOS scales were non-significant.

As the HCR-20 was the only criminogenic need measure that both the HoNOS-S and CANFOR-S were correlated with, a hierarchical regression analyses was conducted to determine where the HoNOS-S accounted for additional variance in the HCR-20 above that of the CANFOR-S. The CANFOR-S was found to explain 59% of the variation in scores on the HCR-20 ($R^2 = .59$, $F(1,56) = 37.327, p < .01$), however the inclusion of the HoNOS-S in the model did not account for any additional variance in the HCR-20 model ($\Delta R^2 = .01$, $F(1,55) = 0.43, p > .05$).

The authors concluded that the CANFOR-S was an adequate forensic mental health needs assessment that incorporates both criminogenic and non-criminogenic needs relevant to individuals with mental illness who have offended. The HoNOS-S on the other hand was associated with specific violence risk factors associated with mental illness rather than broader criminogenic needs. Thus, these two measures contributed different information to care and treatment planning.

Project and Scholarly Work


Summary

This study investigated the interrelationship and overlap between the Camberwell Assessment of Need–Forensic shortened version (CANFOR-S), the Health of the Nations Outcome Scales–Secure (HoNOS-S), the Level Service-Case Management Inventory (LS-CMI), and the HCR-20 in a sample of forensic psychiatric patients residing in a state forensic mental health facility in Australia. Participants of this study were 72 adults (66 males and 6 females), aged between 20 and 62 years of age ($M = 37.78$, $SD = 8.98$) who had committed serious violent offenses (i.e., murder, attempted murder, serious assault). A majority (68%) of the sample had schizophrenia as their primary diagnosis.

The CANFOR-S was used to assess both patient and nurse views of total, met, and unmet needs across 25 domains of individuals experiencing mental illness who have offended. Each domain is coded on a three-point scale from 0 (need is present) to 2 (need is unmet). The HoNOS-S was also used to assess the needs of individuals experiencing mental illness who have offended. It comprises amended versions of the original 12 items HoNOS items and an additional seven-item security scale. Each item is rated on a five-point scale (0-4). A security scale item rated 1 or above indicates that a risk management intervention is needed and for items on the clinical scales, a rating of 2 or above indicates that a care or treatment intervention is needed. The current study, HoNOS-S and CANFOR-S outcome data, routinely completed between patients and their primary nurses, were obtained from the hospital. HCR-20 and LS-CMI data were coded from patient files and clinical documentation by the study authors.

Pearson’s product moment correlations were used to examine the relationship between each of the four measures. The HoNOS-S and CANFOR-S were found to have moderate to strong associations with the HCR-20. CANFOR-S nurse and patient ratings of total needs were significantly and positively correlated with the HoNOS-S clinical and security scales ($r_s 0.57$ and .79), as well as the HCR-20 clinical ($r_s 0.45$ and .68) and risk management scales ($r_s 0.69$ and 0.73). HoNOS-S clinical scales were also significantly, positively correlated with clinical ($r = 0.71$) and risk management scales of the HCR-20 ($r = 0.45$). HoNOS-S security scales were only found to be significantly, positively correlated with the clinical scale of the HCR-20 ($r = 0.67$). The HCR-20 was the only measure found to be significantly correlated with the LS-CMI ($r = 0.75$), correlations between the LS-CMI and the CANFOR or HoNOS scales were non-significant.

As the HCR-20 was the only criminogenic need measure that both the HoNOS-S and CANFOR-S were correlated with, a hierarchical regression analyses was conducted to determine where the HoNOS-S accounted for additional variance in the HCR-20 above that of the CANFOR-S. The CANFOR-S was found to explain 59% of the variation in scores on the HCR-20 ($R^2 = .59$, $F(1,56) = 37.327, p < .01$), however the inclusion of the HoNOS-S in the model did not account for any additional variance in the HCR-20 model ($\Delta R^2 = .01$, $F(1,55) = 0.43, p > .05$).

The authors concluded that the CANFOR-S was an adequate forensic mental health needs assessment that incorporates both criminogenic and non-criminogenic needs relevant to individuals with mental illness who have offended. The HoNOS-S on the other hand was associated with specific violence risk factors associated with mental illness rather than broader criminogenic needs. Thus, these two measures contributed different information to care and treatment planning.

Project and Scholarly Work


Summary

The authors presented on a new treatment model at a secure hospital. In order to facilitate rehabilitation, an extra
security ward was added. Patients needed to meet the inclusion criteria of suffering from a major psychotic disorder and having committed serious violent behaviour towards others to be admitted to the new ward. As their functional level improved, they were transferred to a lower level ward. As part of admittance to this new security ward, a neuropsychological evaluation, the Positive and Negative Syndrome Scale (PANSS) and the HCR-20 were completed. The patients GAF, the HCR-20 and the ADL were administered every 6 months. Biological factors, cortisol and testosterone were also monitored. The treatment process the authors proposed includes a combination of risk assessment and management, structural milieu therapy, progression ladders, anger management, cognitive behavioural therapy, psychosocial support, work therapy and psychotropic medication. The rehabilitation process begins as early as possible and focuses on patient’s level of functioning and coping resources. Through this model, the authors intend to reduce violent recidivism, enhance treatment quality, and prepare for a gradual reduction of the treatment period. It is hoped that the model will inform a future Research Program on the effects of the treatment.

Project and Scholarly Works


Summary

The objective of the current study was to evaluate and promote understanding of the content and implementation of a risk assessment tool within a high security hospital. The authors developed the Structuring Clinical Judgement: Risk (SCJ: Risk) which incorporated all of the items on the HCR-20 plus 6 new subscales (Suicide, Vulnerability, Escape, Risk Scenario Planning, Tilt High Risk Summary, and Risk management Plan). The purpose of the study was to retrospectively evaluate the predictive validity of the SCJ: Risk with regards to institutional violence. Institutional violence was coded on two Levels from hospital incident files 12 months following the initial assessment. The mean age of the sample was 38 years of age. The primary diagnosis was mental retardation (n = 62) but participants also suffered from personality disorder, schizophrenia, and several other Axis I disorders (e.g., pervasive developmental disorder). Participants were mostly Caucasian British (81%). In terms of previous violence, 58% of participants committed a violent index offence, 15% committed one incident exclusive of index, 24% committed between 2-4 violence incidents, and 13% committed 5 or more violence incidents.

During the 12 month follow-up period, 74% of participants were involved in at least one violent incident in the hospital – 57% were involved in a Level 1 incident (physical aggression or any violence resulting in injury) and 68% were involved in a Level 2 incident (general aggression such as verbal aggression or property damage). The mean scores for the HCR-20 are as follows: Total M = 20.87, H scale M = 14.40, C scale M = 4.65, R scale M = 3.61. The mean scores for SCJ: Risk are as follows: Total M = 31.47, H scale M = 25.03, S scale M = 56, V scale M = 1.52, E scale M = .26. The HCR-20 and its subscales produced moderate to large AUCs. The AUC for the HCR-20 total score was significant for any violence (.72), Level 1 (.70) and Level 2 (.76). The H scale was not significant for any of the outcomes. The C scale was significant for any violence (.72), Level 1 (.68) and Level 2 (.77). The R scale was significant for any violence (.66) and Level 1 (.63). The SCI scales were not significant for any outcomes. Only the SCI total was significant for any violence (.68), Level 1 (.66) and Level 2 (.71).


Summary

The current prospective study investigated the predictive validity of the PCL-R, the HCR-20 and the Emotional Problem Scales for institutional aggression in 60 male intellectually disabled (ID) patients. All patients (n = 73) in a high security hospital for England and Wales had previously been assessed in 2003 as part of a wider study.

Participants in the current study were the 60 original participants still present in high security at the 12-month follow-up. The majority of participants were of white British origin (80%). The mean full-scale IQ of the sample was 66.2 (SD = 8.9, range 43 – 76), and the mean age was 38.0 (SD = 8.1) years. The participants had been detained in the high security hospital for an average of 9.0 years (SD = 8.0, range 0 – 35). In total, 81% of the sample had an ICD-10 diagnosis of mental retardation, 54.8% a diagnosis of personality disorder, 28.8% psychotic disorder, and 8% mood disorder (including cases of dual diagnosis).
The PCL-R was completed for all 60 participants by a trained graduate-level psychologist using a combination of file review and interview with a clinical informant (psychiatrist or psychologist). In the wider study, inter-rater reliability was established with a second rater who coded 45 cases using the same methodology. The IRR was good (ICC = .89), although it was slightly lower for cases at the high secure site (ICC = .80). As with the PCL-R, the HCR-20 was completed from a comprehensive file review combined with an interview with a clinical informant. In total, 54 of the 60 patients had a completed HCR-20. Follow-up institutional aggression data were collected from official hospital records. The incidents were divided into several categories: all aggressive incidents; interpersonal physical aggression (labelled Type 1 aggression), and verbal aggression or aggression to property (labelled Type 2 aggression). Then, all incidents were coded into 3 categories that reflected the degree of actual or potential harm to others (low, medium, high). Those incidents in the high category were defined as ‘high risk aggression’ (Type 3).

In total, 76.7% of the sample were involved in at least one aggressive incident during the follow-up: 59.3% had engaged in at least one Type 1 violent incident (Mdn = 1.5, 0 – 103), 70% had engaged in one or more Type 2 incidents (Mdn = 2.5, 0 – 125), and 36% of the sample had engaged in a Type 3 incident (Mdn = 0, 1 – 17). The mean scores for the PCL-R were: Total score M = 18.3, SD = 7.2; Factor 1 M = 7.0, SD = 4.1; Factor 2 M = 9.7, SD = 4.5; 13-item total M = 11.7, SD = 6.0; Items 9, 15, and 17 were prorated). The mean HCR-20 scale score was 22.5 (SD = 4.5).

The PCL-R total score, Factor 1 score, Factor 2 score, and PCL-R 13- item total were not significantly correlated with any type of institutional aggression. By contrast, the HCR-20 total score was significantly correlated with both Type 1 and Type 2 aggression. Neither the PCL-R total, Factor 1, Factor 2 scores, or the 13-item total produced significant AUCs significantly greater than chance for either Type 1 or Type 2 aggression (AUCs = .48 – .59). Both the HCR-20 total score (AUC = .68 – .77) and the EPS externalizing scale (AUC = .72 – .77) significantly predicted both types of aggression. In addition, AUCs obtained for the HCR-20 were significantly greater than those obtained for the PCL-R, except in the case of Factor 2 in relation to Type 1 aggression.


Summary

The current study investigated the predictive utility of the PCL-R and the HCR-20 for 75 male offenders with intellectual disabilities (ID) with respect to positive and negative treatment progress (i.e., moves of patients both within and out of high security). Participants were the entire population of individuals with ID being treated in a high security hospital during 2003. The mean age was 37.0 years (range = 17–68 years). Of the total, 81% had an ICD-10 diagnosis of Mental Retardation, 54% Personality Disorder (specific or mixed), 28.8% Psychotic Disorder, and 8.6% Mood Disorder, with 70% having two or more diagnoses. Participants had already been assessed using a range of measures for the purposes of a wider study addressing ID, personality disorder and risk conducted in 2003. The outcome (positive and negative progress) was coded from institutional records two years after the initial assessment. Active positive progress was defined as movements from the high security facility to a medium security setting. Active negative progress was defined as movements from lower security wards to higher security wards within the hospital, return to prison where the stated reason was lack of suitability for treatment or lack of treatment progress, or moves back to high security from medium security.

Both the PCL-R and the HCR-20 were coded from a full file review plus an interview with a clinical informant. Of the 73 participants, 25 (34.2%) had made active positive progress during the follow-up period and 8 (11%) had made negative progress moves. As the authors predicted, the PCL-R Total score (r = .30), PCL-R Factor 1 (r = .33), PCL-R Total 13 (r = .35), Facet 1 (r = .25) and Facet 2 (r = .36) were all significantly correlated with a negative progress move. However, neither PCL-R Factor 2 nor the HCR-20 Total score were significantly correlated with negative progress. In addition, a positive progress move to medium security conditions was significantly negatively associated with the PCL-R Total score (r = -.36), and with Facet 2 (r = -.30), Facet 4 (r = -.26) of the 4- facet model and the HCR-20 Total score (r = -.32). Only PCL-R Facet 4 (antisocial) was no longer significantly correlated with a positive progress outcome when other variables were controlled.

In the ROC analysis, the PCL-R Total, PCL-R-13 items, PCL-R Factor 1, Facet 1 and Facet 2 were significant predictors of negative progress (AUCs = .80, .82, .84, .77, .85). The PCL-R Total, PCL-R-13 items, PCL-R Factor 1, Facet 2, Facet 4, and the HCR-20 Total score (AUCs = .73, .66, .65, .69, .67,69) all associated with lack of positive progress. With respect to positive progress to medium security conditions, the PCLR-20 demonstrated incremental validity over the HCR-20.

Summary

Risk assessment of individuals with Intellectual Disability (ID) is a growing area of focus. With a sample of 212 forensic patients with ID, the present study assessed the psychometric properties of the HCR-20 within this specific sample. The sample consisted of three subgroups of offenders with ID: a group from a high security unit, a group from a medium/low security unit, and a group from a community forensic unit. In order to establish the convergent validity of the HCR-20, the VRAG, Short Dynamic Risk Scale (SDRS), and Emotional Problems Scale (EPS) were also used. Violence was monitored from nurses’ notes for a one year period.

For the entire sample, the mean scores on the HCR-20 were 19.54, 12.09 ($SD = 4.43$), 4.41($SD = 2.40$), and 3.04 ($SD = 1.78$) for the total score, H, C and R scales, respectively. Interrater reliability was assessed and reported as percent agreement. For the H scale, C scale and R scale, respectively, the interrater reliability was 89.4%, 93.1%, and 82.7%. Group differences were all evident between the different security levels. The lowest security level had significantly lower H scale scores ($F = 59.17$, $p < .001$) and total scores ($F = 30.42$, $p < .001$) than the other security levels. The medium security level had the lowest R scale score ($F = 4.81$, $p < .01$). No group differences were seen on the C scale.

Convergent validity of the HCR-20 was established with the VRAG, SDRS and EPS. Correlations with the VRAG were .65, .23, and .22 for the H, C and R scales, respectively. Correlations with the SDRS were .39, .42, and .29 for the H, C and R scales, respectively. With regard to the EPS subscales, the H scale was significantly correlated with seven of the eleven EPS subscales with coefficients ranging from .25 to .32. The C scale was significantly correlated with nine of the EPS subscales ($r$ ranging from .21 to .39), and the R scale was correlated with four of the EPS subscales ($r$ ranging from .23 to .25). An exploratory factor analysis was conducted on the HCR-20 using principle components analysis with Varimax rotation. A five factor solution was found that accounted for 60% of the total variance.

Significant differences were also found between those who did and did not perpetrate violence in the follow up period. The H scale ($t = -4.01$, $p < .001$), C scale ($t = -3.80$, $p < .001$), and R scale ($t = -2.40$, $p < .02$) all differentiated between those who perpetrated violence and those who did not. Only the H scale differentiated between those who were convicted for a violent offence in the follow up period that those who were not ($t = -6.23$, $p < .001$). The HCR-20 was predictive of violence with AUCs of .72, .68, .67, and .62 for the total score, H, C, and R scales, respectively.

The authors concluded that the HCR-20 is a valuable tool for use with individuals with ID in forensic contexts.

See Also


Project and Scholarly Work

Bauer, P., & Knörnschild, C. (2010, May). *The ignored female minority: Do women have differentiated needs in the forensic setting?* In R. Müller-Isberner (Chair), *Forensic patients with special needs.* Symposium conducted at the annual conference of the International Association of Forensic Mental Health Services, Vancouver, British Columbia, Canada.

Summary

This study examined differences between all men and women admitted to a large psychiatric hospital between 2000 and 2009. Numerous gender differences were found. Women with schizophrenia tended to be older when admitted than men with schizophrenia or other women with personality disorders. Women with schizophrenia tended to more often be married than men with schizophrenia. Women with schizophrenia tended to have lower PCL-R scores than schizophrenic men. Women with personality disorders also tended to have lower PLC-R scores than men with personality disorders.

With respect to the HCR-20, several gender differences were found. Women with schizophrenia had a lower H scale score than men with schizophrenia or women with a personality disorder. Women in general had lower scores on items H2 (young age at first violent incident) and H10 (prior supervision failure) compared to men. Women with schizophrenia also had lower scores than men with schizophrenia and women with a diagnosis of a personality disorder on items H2 (young age at first violent incident),
H4 (employment problems), H5 (substance use problems), and H8 (early maladjustment).

**Project and Scholarly Works**


**Summary**

This was a reliability study of the Swedish translation of the HCR-20 (Belfrage & Fransson, 1997). Six clinicians rated the same 43 patients on the HCR-20 and PCL-R. Over half of the sample (n = 25; 58%) had an index offence of homicide, and the majority (77%) had previous records for criminal offences. The mean age of patients was 40 (range = 24 - 67). The majority of patients had primary ICD-9 diagnoses of mental disorder (70%), 21% were diagnosed as personality disordered, and 9% received other diagnoses. Approximately half of the sample (n = 22; 51%) also had substance abuse diagnoses.

Internal consistency, using Cronbach’s alpha, for the H scale was .96, for the C scale was .89, for the R scale was .85. For the total score, Cronbach’s alpha was .95. Multivariate interrater reliability analyses, using Kendall’s W, produced the following coefficients — Total scale = .81; H scale = .85; C scale = .62; R scale = .56. The HCR-20 correlated with the PCL-R at .64. The Cronbach’s alpha for the PCL-R was .95, and Kendall’s W was .78.


**Summary**

This study evaluated whether Clinical and Risk Management scale scores of the HCR-20 changed over time and whether change predicted future violence. The study sample consisted of 174 forensic psychiatric patients in Sweden. A majority of the sample were male (81%) and had a psychotic disorder (67%). Using a prospective design, the HCR-20 and PCL: SV were administered. C and R scales were measured four times at 6-month intervals. Violence was also recorded between time points.

Repeated measures ANOVA revealed a statistically significant linear decrease on C scores across time (eta² = .12, p < .001). In a hierarchical logistic regression with C (Time 1), C-Change (Time 2-Time1), and their interaction term as predictors, each was significantly predictive of violence that occurred after Time 2 (-2LL = 125.81; Nagelkerke R² = .14; p < .01). A further novel (cluster analytic) approach to uncover differential change within sub-groups of participants indicated that there were distinct groups of patients who differed substantially on C scale

Cross-sectional results showed that the mean scores for the C-scale and the R-scale of the HCR-20 were significantly lower the longer patients had been hospitalized. These results were only significant for the C-scale when comparing the group which had been institutionalized for up to one year against those who had been in for over two years (C-scale p < .038). The R-scale showed significant changes between the one year group and the one-two year group (p = .01) as well as between the one year group and the more than two years group (p < .001).
changes, and violence within these groups changed across time proportionally to change in C scores. While groups significantly differed on PCL: SV Total and F2 scores, overall the PCL: SV was only moderately helpful in separating groups. Risk Management scale results were also presented. The authors concluded that change in dynamic risk factors predicted change in violence, and that there are important differences in degree of change across different groups of patients. Clinical implications were discussed.

**Summary**

This was a retrospective follow-up of 404 forensic patients who had committed violent offences in Sweden, and who were followed up for a period of two years. This study compared to predictive characteristics of the Violence Risk Appraisal Guide (VRAG; Harris, Rice, & Quinsey, 1993) and the H Scale of the HCR-20. The sample was further broken down into two sub-samples or cohorts: 1) 293 violent offenders with ICD-9 diagnoses of personality disorder; 2) 111 violent offenders with diagnoses of schizophrenia.

Across both groups, the AUC of the ROC for the H Scale was .71 (95% CI = .66 - .76). At the cut-off score of 12 on the H Scale (the inflexion point), sensitivity = .71; specificity = .61; positive predictive power = .35, and negative predictive power = .88. For the VRAG, the AUC was .68 (95% CI = .63 - .73). At the cut-off score of 13 on the VRAG (the inflexion point), sensitivity = .50; specificity = .77; positive predictive power = .39, and negative predictive power = .84.

In the personality disordered cohort, the AUC of the ROC for the H Scale was .71 (95% CI = .66 - .76). At the cut-off score of 12 on the H Scale (the inflexion point), sensitivity = .72; specificity = .60; positive predictive power = .38, and negative predictive power = .86. For the VRAG, the AUC was .68 (95% CI = .62 - .73). At the cut-off score of 13 on the VRAG (the inflexion point), sensitivity = .57; specificity = .71; positive predictive power = .40, and negative predictive power = .83.

In the schizophrenia cohort, the AUC of the ROC for the H Scale was .66 (95% CI = .56 - .75). At the cut-off score of 8 on the H Scale (the inflexion point), sensitivity = .88; specificity = .36; positive predictive power = .19, and negative predictive power = .95. For the VRAG, the AUC was .60 (95% CI = .50 - .69). At the cut-off score of 0 on the VRAG (the inflexion point), sensitivity = .68; specificity = .53; positive predictive power = .20, and negative predictive power = .91.

Grann et al. concluded that both the H Scale and the VRAG predicted violence significantly better than chance (except for the VRAG in the schizophrenia group). They comment that the obtained values could under-represent the actual predictive accuracy of the instruments because several items on each scale had to be “approximated.” The sensitivity of the H Scale tended to be greater than that for the VRAG, whereas the specificity of the VRAG tended to be greater. Among the schizophrenia group, only the H Scale was better than chance.

**Summary**

Using data from Grann et al. (2000), this investigation evaluated the relative accuracies of different options for weighting H scale scores. The authors used five approaches: nonweighted, Nuffield approach, logistic regression model (one-by-one), logistic regression model (11-term algorithm), and artificial neural network. They split the sample into training (or calibration) and validation seeds or subsets. Results showed that the unweighted procedure produced the largest average AUC value (.72), compared to the Nuffield approach (.71), logistic regression one-by-one (.71), logistic regression 11-term algorithm (.68) and artificial neural network (.64). These findings are consistent with research showing that unit-weighted predictors are often as accurate as optimally-weighted procedures.

**Summary**

The retrospective study examined if patients admitted to forensic psychiatric care could decrease their risk for violence over time and, if so, which factors played an important role in contributing to their decrease in aggressive behavior. The sample used in this study consisted of 267 individuals admitted to a maximum-security forensic psychiatric clinic in Sweden between 1997 and 2010. The average age of the participants at time of admission was 35 years (SD = 10.8). A majority (80%) of the sample was male, had a substance abuse problem (65%), and a psychotic disorder (63%). A total of 233 (87%) patients were assessed.
using the PCL: SV (M = 11.8, SD = 5.3). Twelve percent of patients had scores of 18 or higher.

The HCR-20 was coded from comprehensive file reviews of information which had been routinely collected by the clinic staff. Assessments at admission were compared with a second, and most recent, risk assessment. The mean time between admission and the patient's initial assessment, the first and second assessment, and the first and most recent assessment were 11, 9, and 43 months, respectively. Decreased risk for violence was operationalized as a reduction in scores on either the C or R scale of the HCR-20.

The authors found that total score on both the C and R scales showed a significant reduction, both over a short period of time as well as a longer period of time. Scores on thirteen out of 15 risk factors significantly decreased after more than 2 years of treatment. A comparison of the risk assessment on the C scale reveals that between assessment 1 and 2, the total score dropped significantly from 5.82 (SD = 2.15) to 5.50 (SD = 2.15). In terms of item level, reductions were seen in all items except for C4 (impulsivity). When looking at the comparison of the first and most recent assessments, the total score dropped even more, 5.0 (SD = 2.16), also the scores on all items decreased. The total score on the R scale decreased and showed the greatest reduction between the first assessment 5.08 (SD = 1.70) and the second assessment 4.77 (SD = 1.81). Regarding item level, R1 (plans lack feasibility), and R4 (noncompliance with remediation), showed a decrease, while no decrease was found for R2 (exposure of destabilizers), R3 (lack of personal support), and R5 (stress) when comparing the first and the second assessment. When comparing the first assessment to the most recent assessment, a risk reduction was found also for R2 (exposure of destabilizers), but R3 (lack of personal support), and R5 (stress), still did not show a significant change in risk.


**Summary**

Dyslexia does not cause criminal behavior, but it may worsen aggressive behavior tendencies. In this study, aggressive behavior and risk of future violence were compared between forensic psychiatric patients with (n = 18) and without dyslexia (n = 14). The study sample consisted of 32 (26 male and 6 female) forensic psychiatric patients from a high-security forensic hospital in Sweden. Of the sample, 65.63% had been sentenced for violent crimes, 12.5% for sexual crimes, and 3.12% for other crimes.

Dyslexia was assessed using the Swedish phonological processing battery ‘The Pigeon’. The HCR-20 was completed by trained assessors based on interviews with the patients and their caregivers, forensic psychiatric investigators, sentences, and journals. Aggression was determined using Swedish version of the patient self-report Aggression Questionnaire which is divided into four subscales: physical aggression, verbal aggression, anger, and hostility. Total score on the Aggression Questionnaire was found to be positively related to total score of the HCR-20 (r = 0.54, p < .01). Total score on The Pigeon was not significantly related to Aggression or HCR-20 scores, the only exception being Anger (r = -0.35, p < .05). Anger, in turn was found to explain 35% of the variance in total score of the HCR-20 (β = .59, p < .001)

Patients with dyslexia (M = 80.11, SD = 15.15) self-reported significantly more aggressive behavior than did patients without dyslexia (M = 62.86, SD = 15.97). For the dyslexic group, mean scores were 15.11 (SD = 3.41), 6.06 (SD = 1.51), 5.50 (SD = 2.01) and 26.67 (SD = 5.05) on the historical, clinical, risk management and total scales of the HCR-20, respectively. For the non-dyslexic group, mean scores were 13.57 (SD = 3.46), 5.21 (SD = 2.46), 4.64 (SD = 1.60) and 23.43 (SD = 4.88) the historical, clinical, risk management and total scales of the HCR-20, respectively. While there was a tendency for patients with dyslexia to receive higher scores on the HCR-20 compared to patients without dyslexia, this was non-significant. The authors note that a follow-up study should be conducted using the HCR-20 to examine whether patients with dyslexia relapse more often in violent crimes than patients without dyslexia


**Summary**

The purpose of this study was to compare the scores on the HCR-20 between male and female forensic patients. Using the official Swedish translation of the HCR-20, all female patients (n = 63) who entered a Swedish forensic facility over 10 years were assessed with file, and, where possible, also with interview. Comparisons were made with all 85 male patients admitted to two Swedish forensic hospitals in 1998.

The female sample was younger (30.8 vs. 35.1 years), more often diagnosed with a personality disorder (55.6% vs. 36.5%, specifically borderline [85.7% vs. 25.8%], and less...
often antisocial [0.0% vs. 25.8%]). Females were less often admitted after committing violent crimes (9.5% vs. 31.8% murder; 17.5% vs. 31.8% other violent crimes), and more often admitted from general psychiatry due to violence (42.9% vs. 2.4%).

There were no differences in scale or total scores between genders. Total score = 24.76 (SD = 6.95) female, 25.51 (SD = 7.92) male; H scale = 12.94 (SD = 3.58) female, 13.81 (SD = 4.21) male; C scale = 5.11 (SD = 2.57) female, 5.00 (SD = 2.48) male, R scale = 6.71 (SD = 2.85) female, 6.68 (SD = 2.80) male.

There were differences on some of the items, likely reflecting the general differences between genders. Males scored higher on Previous Violence (H1), Young Age... (H2), Substance Use Problems (H5), and Negative Attitudes (C2). Females scored higher on Personality Disorder (H9), Impulsivity (C4), and Stress (R5).


Summary

The Swedish version of the HCR-20 was coded on 40 male forensic psychiatric patients in a postdictive study of the HCR-20 and PCL:SV. There were 22 recidivistic patients and 18 non-recidivistic patients who were matched on demographic, clinical, and criminal variables. The rater was blind to recidivism status. Overall, the recidivistic group scored 8 points higher than the nonrecidivistic group (Ms and SDs = 30.77 [7.22]; 22.39 [6.85], respectively). Although not reported, this represents a Cohen’s d of 1.19, which is a large effect size. All persons (n = 11/40) with scores above 34 on the HCR-20 recidivated. Interestingly, Strand et al. report that for patients who scored between 24 and 28, prediction was random. However, all recidivistic patients in this scoring range scored 2/2 on R5 (Stress), and this item alone differentiated the two groups (for this scoring range).

The area under the curve of the receiver operating characteristic analysis was .80 for the HCR-20, and .70 for the PCL:SV. Using a cut-off score of 29/40 on the HCR-20, sensitivity was reported to be .89 and specificity .64. With a cut-off of 17/24 on the PCL:SV, sensitivity was .89, and specificity was .59.

Surprisingly perhaps, the items from the Clinical and Risk Management scales were much stronger in separating the two groups than was the Historical scale. Strand et al. point out that this finding may stem from the fact that the patients in their sample, given their offences and dispositions to a forensic hospital, were homogenous on historical factors.


Summary

This was a long-term predictive validity assessment (with retrospective data collection) of the VRAG and the Historical part of the HCR-20. The sample consisted of 106 violent offenders with schizophrenia in Sweden. The mean detention time for the offenders was 18 months. All subjects were followed from discharge or start of probation until each subject was at least 5 year from that point. The average time from discharge at follow-up was 86 months. The definition used for this study of violent recidivism was closely matched to the definition adopted in the VRAG calibration sample; a reconviction of attempted or completed homicide, assault, all sex crimes, armed Robbery and forcible confinement. During the follow-up period, 29% of the sample was reconvicted of a violent crime.

The following H-scale items were significantly correlated with violent recidivism: H7 (r = .42; p < .01), H1 (r = .36; p < .01), H8 (r = .20; p < .05), H9 (r = .28; p < .05), H2 (r = .24; p < .05), H10 (r = .38; p < .01), and H5 (r = .3; p < .01). The AUC for the whole H-scale to predict violent recidivism was .76, while the AUC for the VRAG was .68.

The predictive validity of both the VRAG and the H-scale was considered to be moderate. There was a reported trend for the H-scale to perform slightly better compared to the VRAG. For the H-scale, most of the items had a positive correlation to recidivism and contributed well to the overall performance of the sub-scale. When similar items from the VRAG and the H-scale were compared several differences emerged.

See Also


Project and Scholarly Work

**Abridged abstract (English translation of the study not available):**

**Background:** Although Dutch forensic psychiatry is making increasing use of structural risk assessment scales, the controversy about the value and usefulness of these instruments continues unabated. **Aim:** To provide an overview of the psychometric qualities of the instruments used most often in the Netherlands for risk assessment in adults. **Method:** Dutch data about the Historical, Clinical, and Risk Management (HCR-20), the Sexual Violence Risk-20 (SVR-20), the Psychopathy Checklist-Revised (PCL-R) and two Dutch instruments, the ‘Historische, Klinische enToekomstige Risico-indicatoren-30’ (HKT-30) and the ‘Forensisch Psychiatrische Profielen’ (FP-40) were reviewed. In addition, data relating to the unstructured clinical judgment were studied. **Results:** The inter-rater reliability values of the instruments discussed were in general satisfactory, but the internal consistency was often unsatisfactory. Except in some studies, the predictive validity was in general reasonable. **Conclusion:** At present, caution is called for with regard to the assessment of the risk of recidivism when this is based purely on risk assessment scales or purely on the unstructured judgment. Perhaps it is simply not possible to predict recidivism more accurately. Until there are some new developments in this area, it seems advisable to combine as many data as possible about a person under investigation derived from assessment scales and clinical judgment and to compare the outcome with the conclusions of the other professionals.

**Project and Scholarly Work**


**Summary**

In this prospective study preliminary results are reported from a multicenter randomized clinical trial on the effectiveness of Schema Therapy (ST) for hospitalized TBS patients (n = 30) with Antisocial, Borderline, Narcissistic, or Paranoid Personality Disorder. Patients at seven TBS clinics were randomly assigned to received three years of either ST or treatment as usual (TAU) and were assessed on several outcome variables, such as recidivism risk (assessed every six months using the HCR-20, SVR-20, and START), personality disorder symptoms (SID-IV, SNAP-I), and successful re-integration into the community. Of the sample, a majority was of Dutch origin (90%), diagnosed with Antisocial Personality Disorder (86.7%) and had received a TBS sentences for committing a violent offense (90%). Approximately 26.7% of the sample had PCL-R scores of 30 or higher.

Inter-rater reliability was determined using a subsample of 16 cases. ICC values were .81 and 1.0 for the HCR-20 overall judgment of risk level within the hospital and ratings of risk outside the hospital, respectively. The inter-rater reliability for the PCL-R total score was .88.

Fisher’s exact test and Cox regression survival analysis with PCL-R scores as a covariate were used to compare the two treatment conditions with respect to the number of days required to obtain permission for supervised leave and unsupervised leave, respectively. In both analyses, results suggest that the ST patients received leave more rapidly than the TAU patients, though these differences were not statistically significant.

The authors also conducted repeated measures ANOVA to analyze the effect of ST versus TAU on HCR-20 scores over the course of treatment, using centered PCL-R scores as a covariate. Results indicate that while HCR-20 scores improved more rapidly in ST patients compared to TAU patients, no statistically significant effects of treatment were found. When centered PCL-R scores were entered as a covariate, a highly significant effect of PCL-R scores on patients’ HCR-20 scores was found. There was also a statistically significant effect of time, but again no main effect of treatment condition on HCR-20 scores. Other outcome variables or interactions of treatment conditions with PCL-R scores were not examined because of the low statistical power in the sample, which could have impacted the main effect analyses as well.

**Project and Scholarly Work**


**Summary**
This study evaluated the utility of the HCR-20, PCL-R, and VRAG in predicting negative outcomes of people found NCRMD. The sample comprised 172 insanity acquittees (20 women and 152 men) appearing before a criminal Review Board in British Columbia. Participants’ mean age was 34.17 years (SD = 9.70). Most (91%) had primary diagnoses of a psychotic disorder (6% organic mental disorder; 2% anxiety or other disorder; 1% substance abuse disorder). Almost half (42%) had secondary diagnoses of substance abuse or dependence and 31% were diagnosed with a personality disorder.

The C and R scales were completed by psychiatrists as part of their routine assessment prior to the Review Board Hearing. The H scale was completed by a research assistant using file material. For 67 participants, the PCL-R was completed using both an interview and file material; for the remainder of the sample, only file material was used. All instruments were completed prospectively except for the VRAG, which was coded at the end of the study period.

The mean PCL-R scores were: total = 16.51 (SD = 7.27); F1 = 5.92 (SD = 3.14); F2 = 8.84 (SD = 4.00). The mean VRAG score was 0.58 (SD = 8.92). The mean HCR-20 scores were: total = 22.67 (SD = 6.53); H scale = 12.80 (SD = 3.64); C scale = 4.47 (SD = 2.54); R scale = 5.41 (SD = 2.71).

Elopers had significantly higher scores on the H scale compared to non-elopers; t (109.62) = 3.58, p = .001. The two groups did not have significantly different scores on the C (p = .548) and R (p = .342) scales. Elopers also had significantly higher PCL-R scores and were placed into higher VRAG score bins relative to non-elopers. Cox regression analyses using the HCR-20 scales, with time at risk calculated as time spent in hospital during the study period, yielded the same pattern of results as the univariate analyses. When the individual HCR-20 items were examined, only Item H10 (prior supervision failure) was positively and significantly associated with risk of elopement. When HCR-20 scales were compared to PCL-R total scores and VRAG bin scores controlling for age, none of the variables was related significantly to risk of elopement.

There were 109 participants who were released on conditional discharge. Compared to participants not released, those who were released had significantly lower mean scores on the C scale (t = 6.74, p \leq .001) and R scale (t = 9.61, p \leq .001). The groups did not have significantly different H scale (p = .843), PCL-R (p = .603), or VRAG bin scores (p = .790) Cox regression analyses indicated that R scale scores were associated with likelihood of release (Wald = 23.06, p \leq .001), but H (Wald = .42, p = .517) and C scale (Wald = .36, p = .550) scores were not. When individual HCR-20 items were considered, negative and significant associations with release were found for previous violence, active symptoms of major mental illness, and plans lack feasibility. In another Cox regression analysis that compared the three HCR-20 scales, PCL-R, and VRAG bin scores controlling for age, a significant (positive) relation was found only for the R scale. Age was associated negatively with release.

Of the 109 participants released, 43 were returned and one committed suicide. The following analyses consider only the first rehospitalization in those cases were there were multiple returns for the same individual. Neither univariate analyses nor Cox regression analyses indicated significant differences on any of the measures between those who were successful or failed on release. When the individual HCR-20 items were considered, a positive and significant relationship was found only for active symptoms of major mental illness. When the dependent variable was narrowed to rehospitalization following a significant security problem in the community, PCL-R (Wald = 9.41; p = .002) and R (Wald = 3.89; p = .049) scale scores were significant positive predictors. The H scale was significantly but negatively related to this return following a security problem (Wald = 6.89, p = .009).

**Project and Scholarly Work**


**Abridged abstract**

**Objectives**: Professional risk assessment provides estimates of future risk in terms of nature, frequency and severity as well as may determine treatment services required of an offender under parole or probation supervision. This study examined the predictive validity of the HCR-20 in a Lithuanian forensic context. **Methods**: This research includes a sample of criminal offenders and a sample of forensic psychiatric patients (a total of 118 participants). The HCR-20 was coded on the basis of file information. **Results**: The mean HCR-20 scores were: Total score = 14.96 (SD = 6.56), H-scale = 8.42 (SD = 3.61), C-scale = 2.57 (SD = 2.16), R-scale = 3.89 (SD = 2.55). ROC analyses of the HCR-20 subscales showed AUC’s of .72 for the H Scale, .69 for the C Scale and .58 for the R Scale. **Conclusions**: The HCR-20 total score and final risk judgments were significantly more accurate in predicting violent recidivism (p < .05). Results indicate the predictive ability of the instrument may be maximized when judgments of final risk are used rather than an actuarial approach wherein individual risk factors are summed.
Project and Scholarly Works


Summary

This prospective study compared the predictive validity of the Brøset Violence Checklist (BVC), the Dynamic Appraisal of Situational Aggression (DASA), and the HCR-20 Clinical scale for imminent inpatient aggression over a 24-hour period. Aggressive behaviors were recorded on a modified version of the Overt Aggression Scale (OAS). Acts of aggression were classified as either verbal threats (threats to kill or cause bodily harm to others) or interpersonal violence (biting, hitting, kicking, punching, and throwing objects intending to injure).

The study sample consisted of 70 patients (55 males and 15 females) who were present at the start of the study period (June 2002) or admitted into the acute units of a high-security forensic mental health hospital in Australia. A majority of the sample was Caucasian (78.6%), had psychotic disorders (80%) or substance abuse/dependence (74.3%) and committed a violent index offense prior to their admission to the hospital (71.4%). A total of 90 incidents of inpatient aggression over the 24-hour period were recorded: 46.7% of the sample was involved interpersonal violence and 81.1% made verbal threats.

Unit nursing staff completed the BVC, DASA, and HCR-20 Clinical scale for each patient. Although inter-rater reliability analyses were not conducted, the second and third author trained the nurses in the use and rating of the tools. The authors note that because the DASA has an overlap of two items with the BVC and an overlap of another two items with the HCR-20 Clinical scale, these overlap items were not rated again and were incorporated into the tabulation of the total score for the DASA. Because of the similar items between scales, it is unsurprising that the three measures’ total scores were significantly correlated with each other. Correlations were .73 between the HCR-20 C scale and the DASA, .43 between the HCR-20 C scale and the BVC, and .62 between the BVC and DASA. All correlations were significant.

Results showed that the DASA and BVC had large effect sizes with violence, and that the C scale had a moderate effect size. The DASA and BVC were significantly more accurate than the HCR-20 Clinical scale for predicting interpersonal violence, verbal threat, and any inpatient aggression (i.e., presence of interpersonal violence and/or verbal threat). In regards to any inpatient aggression, AUCs were .68, .76 and .77 for the HCR-20 C Scale, DASA and BVC, respectively. In regards to interpersonal violence, AUCs were .72, .83 and .75 for the HCR-20 C Scale, DASA and BVC, respectively. In regards to verbal threat, AUCs were .68, .77 and .77, for the HCR-20 C Scale, DASA and BVC, respectively.

Overall, these findings support the use of the DASA, BVC, and HCR-20 Clinical Scale for predicting imminent aggression within inpatient psychiatric settings. Notwithstanding that the BVC and the DASA generally performed better than the HCR-20 Clinical scale, the present study showed that all three measures had modest to excellent predictive validity for inpatient aggression during a 24-h follow-up. The authors noted that inter-rater reliability indices were not obtained for the various measures in this study and this limitation should be addressed in future research.


Summary

This study compared the predictive accuracy of dynamic risk assessment tools with static measures for violence over the short term (up to 1 month) and medium term (between 1 and 6 months) in a forensic psychiatric inpatient setting. The study sample consisted of 66 patients who were present or admitted into an acute ward of a high-security forensic mental health hospital in Australia. A majority of the sample was male (80.3%), Caucasian (78.8%), admitted as security patients (i.e., prison transfers, 81.4%) and had a violent index offense (71.2%) prior to admission. Of the sample, 84.8% presented with psychotic disorders during their admission to the hospital, and 19.7% also presented with personality disorders.

The HCR-20, START, LSI-R: SV, PCL-R, and VRAG were retrospectively coded from patient case file materials by the study’s first author who was blind to inpatient aggression. In-patient aggression was assessed over a 6-month period and was categorized into: interpersonal violence (which included biting, hitting, kicking, punching and throwing objects intending to injure), verbal threat (which included threats to kill or cause bodily harm to others), and any inpatient aggression (which included both interpersonal violence and verbal threat). Of the sample, 33.3% were violent towards staff, 24.2% were violent towards other patients, 15.2% made verbal threats of physical harm to others, and 13.6% engaged in property damage.
AUC values were reported for each measure and interpersonal violence outcome at 1, 3, and 6 months. To control for Type I error, false discovery rate (FDR) corrections were also conducted. Results indicated that HCR-20 total scores demonstrated the highest predictive accuracy (AUC = .78) for interpersonal violence over 1 month, and the HCR-20 Clinical and Risk Management scales, the PCL-R total score, and the START Risk scale also significantly predicted interpersonal violence at 1-month follow-up after FDR corrections (AUCs = .73, .75, .72, and .75, respectively). In addition, the HCR-20 Total score, Clinical and Risk Management scales, and START Vulnerability scale significantly predicted interpersonal violence at 3-month follow-up after FDR corrections (AUCs = .75, .75, .75, and .79, respectively). There were no measures that significantly predicted interpersonal violence at 6-month follow-up after FDR correction, even though the PCL-R Facet 3 and the START Vulnerability scale showed acceptable levels predictive accuracy (AUCs = .71 and .74, respectively). Although the HCR-20 Clinical and Risk scales, PCL-R Facet 2, as well as the START Risk scale showed moderate to strong predictive accuracy for verbal threat (AUCs = .72-.84), none of the measures was significantly predicted verbal threat during the follow-up periods after FDR corrections.

Overall, the START Vulnerability scale appeared to be the most predictive of any inpatient aggression (i.e., interpersonal violence or verbal threat) both over the short term (i.e., 1 month; AUC = .74) and the medium term (i.e., 3 and 6 months; AUCs = .83 and .74, respectively), though it only significantly predicted any inpatient aggression during 1- and 3-month follow-ups after FDR corrections. In addition, the HCR-20 Total, Clinical, and Risk Management Scales significantly predicted any inpatient aggression during 1-month (AUCs = .72-.73) and 3-month (AUCs = .76-.78) follow-ups after FDR corrections. The PCL-R total score also significantly predicted any inpatient aggression during 3-month follow-up (AUC = .72) after FDR correction. However, the HCR-20 Historical scale, the LSI-R: SV, and the VRAG were generally inadequate for predicting any inpatient aggression. The authors did not report whether differences in predictive accuracy of risk instruments was statistically significant.

Overall, results showed that dynamic measures were generally more accurate than static measures for short- to medium-term predictions of inpatient aggression. Most of the dynamic risk assessment measures significantly predicted inpatient aggression during 1-month (i.e., short term) and 3-month (i.e., medium term) follow-ups. In particular, the Clinical and Risk Management Scales of the HCR-20 were largely responsible for its predictive accuracy in the short to medium term; the Historical Scale performed inadequately in this context. The HCR-20 Historical Scale did not significantly predict any type of inpatient aggression and showed modest predictive accuracy throughout the follow-up periods. Overall, findings support the assumption that risk assessment measures that consists of dynamic or clinically relevant variables are likely to play important roles in predicting violence in the short term.


Summary

The present study was based conceptually on the premise that continuous monitoring of patients is needed in order to assist daily decision making of ward staff, and that self-harm and violence are highly related. From this backing, the present study investigated the predictive validity of the Dynamic Appraisal of Situational Aggression (DASA) and the HCR-20 Clinical scale in predicting imminent violence and self harm. For the HCR-20, the usually scoring scheme was amended and each item was scored simply as present or absent.

On a high secure forensic psychiatric unit that houses high risk personality disordered offenders, the DASA and C scale were rated daily by nursing staff. Aggression and self harm were also assessed on a daily basis. Both of these instruments were found to be predictive of aggression and self harm. With regards to aggression, the DASA yielded a slightly larger effect compared to the C scale (AUCs of .65 compared to .63). The same pattern was true of prediction of self harm (AUC of .67 for the DASA and .66 for the C scale). The authors concluded that acts of aggression and self harm can both be predicted on a daily (imminent) basis using these instruments.

See Also


Project and Scholarly Work


HCR-20 Review and Annotated Bibliography
management) in a Belgian forensic population. Poster presented at the annual conference of the International Association of Forensic Mental Health Services, Munich, Germany.

Summary

This study reported on the descriptive statistics of the HCR-20 as well as assessing the relations between the HCR-20, the PCL-R, and the Buss and Perry Aggression Questionnaire (Buss and Perry, 1992). This study used 86 French speaking male adult offenders confined in a Belgian forensic hospital.

Types of offenses measured were: violent offenses, non-violent offenses and any sex offenses. Common items between the HCR-20 and the PCL-R and the AQ were omitted. The omitted items were H7 (psychopathy), C1 (intoxication) and H1 (past violent behaviour) from the HCR-20.

The HCR-20 total score had adequate inter-rater reliability ($r = .73$) and good internal consistency (Cronbach’s alpha = .74). The inter-rater reliability for the H-scale alone was ($r = .85; p < .01$) with an internal consistency alpha of .61. The inter-rater reliability for the C-scale alone was ($r = .65; p < .05$) with an internal consistency alpha of .47. The inter-rater reliability for the R-scale alone was ($r = .64; p < .05$) with an internal consistency alpha of .54.

The HCR-20 and the PCL-R were highly correlated across most of their scales. The total, H and C scales from the HCR-20 were all significantly ($p < .01$) and highly ($r$’s $> .4$) correlated with the PCL-R total, Factor 1 and Factor 2 scales. The HCR-20 R scale was only correlated at the $p < .05$ level and with $r$’s between .22 and .25 with the PCL-R scales. Using only a sub sample of 70 men, the HCR-20 scales showed far fewer significant correlations with the AQ. The HCR-20 total score ($r = .3; p < .05$) and the H-scale score ($r = .39; p < .01$) were correlated with the AQ total score. The HCR-20 total score ($r = .34; p < .01$) and the H-scale score ($r = .46; p < .01$) were also correlated with the AQ physical aggression score. The other HCR-20 scales were not significantly related to the AQ scores. The HCR-20 scores were correlated to a few types of violent offenses. The HCR-20 total score was correlated with violent theft ($r = .26; p < .05$) and with assault and battery ($r = .3; p < .01$). The H scale was also correlated with violent theft ($r = .26; p < .05$) and with assault and battery ($r = .37; p < .01$). The C scale was correlated with kidnapping ($r = .26; p < .05$).

The HCR-20 scores were correlated to a few types of non-violent offenses. The HCR-20 total score was correlated with theft ($r = .28; p < .01$). The H scale was also correlated with theft ($r = .27; p < .05$) as well as drug offenses ($r = .24; p < .05$).

Lastly, the HCR-20 scores were correlated with indices of homicide. The HCR-20 total score was correlated with psychotic homicide ($r = -.74; p < .01$). The H scale was also correlated with psychotic homicide ($r = -.67; p < .05$). The C scale was also correlated with psychotic homicide ($r = -.64; p < .05$) and reactive homicide ($r = -.56; p < .05$) and with instrumental homicide ($r = .71; p < .01$).


Summary

This research was a mixed time perspective study using the HCR-20, PCL-R and the Violence Risk Appraisal Guide (VRAG; Harris, Rice, & Quinsey, 1993). This study used 80 subjects to measure correlations of the above measures with type of offense, and a subset of 58 subjects to conduct Kaplan-Meier survival analyses and ROCs. Type of offense fall under two categories: general recidivism (any offense committed after release) and violent recidivism (homicide, assault and battery, theft with violence, any sex offense). The mean follow up period after release was 994 days.

In terms of general recidivism, the PCL-R was correlated ($r = .26; p < .05$) with drug offenses, ($r = .33; p < .01$) with carrying a weapon, and ($r = .46; p < .01$) with theft. The HCR-20 was correlated ($r = .24; p < .05$) with drug offenses, ($r = .23; p < .05$) with carrying a weapon, and ($r = .40; p < .01$) with theft. The VRAG was correlated ($r = .26; p < .05$) with drug offenses and ($r = .47; p < .01$) with theft.

In terms of violent recidivism, the PCL-R was correlated ($r = .39; p < .01$) with assault and battery and ($r = .48; p < .01$) with violent theft. The HCR-20 was only correlated ($r = .32; p < .01$) with violent theft. The VRAG was correlated ($r = .29; p < .05$) with assault and battery and ($r = .38; p < .01$) with violent theft.

In terms of predicting general recidivism, the PCL-R had an AUC of .78. The VRAG had an AUC of .86 and the HCR-20 had an AUC of .79. With predicting violent recidivism, the PCL-R had an AUC of .85, the VRAG had an AUC of .84, and the HCR-20 had an AUC of .78.

The following Pearson correlations between the measures were found: PCL-R was correlated with the VRAG ($r = .67$) and the HCR-20 ($r = .83$), while the VRAG was correlated with the HCR-20 ($r = .68$).
Project and Scholarly Works


Summary

The objective of the present study was to determine whether risk assessments are used at the same frequency in cases of NCRMD (not criminal responsible by reason of a mental disorder) as they are in other legal circumstances. Between October 2004 and August 2006, 96 men were assessed using the HCR-20 prior to their Review Board hearings. The inter-rater reliability was excellent (ICC = .87). In addition, the authors reviewed information presented by clinical psychiatrists at the disposition hearings, coding for factors from the HCR-20 that were considered to be associated with violence.

The analysis was based on the kappa between the factors identified by the research team and the factors mentioned by psychiatrist in his or her report to the Review Board, the factors discussed in the hearing, and those the Review Board considered in their decision. All the items of the HCR-20 were dichotomized on the basis of the absence (0) or presence (1 or 2) of the item. Very few of the risk factors the research team considered as potentially relevant were actually mentioned during the hearing process. Exceptions to this finding were “prior violence” and the presence of “serious mental disorders”. For the H subscale, there was little or no agreement for the majority of items; “substance use problems” had moderate agreement. Although the agreement for the C subscale was better, only two factors had moderate agreement: presence of “active symptoms of mental illness”, and “resistance to treatment”. None of the R subscale items exhibited moderate or better agreement. The results were essentially identical even when the authors considered only forensic clinicians. However, forensic clinicians were more preoccupied with substance abuse problems (kappa = .72) and this had implications for the justification of decisions (kappa = .68). Overall, agreement on personality disorder and psychopathy was weak but agreement among forensic clinicians was very low, comparable to that observed among general psychiatrists. The authors concluded that, overall, there is little application of empirically supported risk assessments. But, the results could be biased given that the role of any expert is to provide an opinion, not necessary to justify the opinion which might have explained the lack of risk assessment information. In addition, patients are usually known to the Review Board and hence some information may have been omitted.


Summary

This study examined the extent to which indicators from the HCR-20 were represented in clinical opinions about people declared Not Criminally Responsible on account of Mental Disorder (NCRMD) at annual review board hearings. The authors’ aims were not to evaluate the use of the HCR-20, but rather to examine the extent to which clinical testimony and review board decisions reflected HCR-20 items which have been empirically validated as risk factors. Participants (n = 96, 100% male) were recruited from a forensic psychiatric hospital and two large civil psychiatric hospitals in Eastern Canada designated for the care of people declared NCRMD.

The risk assessments presented by clinicians at annual review board hearings and the boards’ rationale for the release or detention of people found NCRMD were contrasted with risk assessments conducted by the research team using the HCR-20. Each HCR-20 item was scored dichotomously as present or absent. An item not mentioned in the psychiatrist’s report, at the hearing, or in the reasons cited for the board’s decisions was considered absent. Analysis was based on the coefficient of agreement (kappa) between scale items identified by the research team by the interviews and file reviews, and conversely, the items recorded by clinicians in the report to the review board, the indicators discussed during the hearing, and those cited by the board in its decision.

Results indicate that few of the risk factors identified on the HCR-20 were mentioned during the hearing process, whether in the clinical reports, discussions at the hearing, or in the reasons cited for the disposition. The only noteworthy agreement between potentially relevant and actually mentioned risk factors were for H1 (previous violence) and H6 (presence of a major mental illness). With the exception of these 2 items, which had perfect agreement, and substance use problems, for which agreement was considered moderate (0.41), there was very little or no agreement for historical items of the HCR-20. With regards to clinical items, moderate levels of agreement were found for only 2 indicators reported by clinicians: C3 (active symptoms of major mental illness) and C5 (unresponsive to treatment). None of the risk management items yielded agreement higher than the moderate range.

Crocker, A. G., & Côté, G. (2009b). Evolving systems of care: Individuals found not criminally responsible on account of mental disorder in custody of civil and
Summary

The present study was undertaken to describe the psychological, sociological, criminogenic, and risk of individuals found NRMD and subsequently remanded to either civil psychiatric hospitals (CPH) or a forensic psychiatric hospital (FPH). Participants were recruited between October 2004 and August 2006 from two CPHs and one FPH prior to their review board hearings. The final sample consisted of 96 men between the ages of 18 and 65, 60 from the FPH and 36 from the CPHs. The participants’ mean age was 39.02, with patients in the FPH being slightly older. The majority of them had not completed high school (62.1%). Of the total sample, 87.5% had a history of prior psychiatric hospitalizations, and 84% had a diagnosis of schizophrenia. No major differences were found between the two settings in terms of the participants’ Axis I diagnoses, personality disorders, PCL-R scores, VRAG scores, or type of index offence, with the exception that individuals accused of homicide were sent to the FPH.

Inter-rater reliability was assessed based on 10% of the cases. Intraclass correlation coefficients revealed excellent reliability for the HCR-20 (0.87) and the PCL-R (0.95). Participants remanded to the CPH had a slightly higher HCR-20 total score (M = 23.69, SD = 5.94) than those remanded to the FPH (M = 21.97, SD = 5.60), a small effect size (d = .30) between groups. The two groups, CPH and FPH, had similar scores on the H (M = 14.36 and 13.75, respectively) and C scales (M = 4.67 and 4.65, respectively). The R scale scores significantly differed between the groups, with men remanded to CPH having higher scores (M = 4.67) compared to FPH (M = 3.54). This group difference was mainly accounted for by differences on two items: R2 (destabilizing factors) and R4 (non-compliance with remediation attempts). No differences were found between individuals remanded for a violent offence compared to a non-violent offence on the HCR-20 (total or subscale scores).

The authors discussed that FPH may be more suited to and capable in dealing with individual’s risk management needs, which results in the lower R scale scores in this setting. They also discussed implications of the lack of major differences between the two groups, as they believe that men remand to FPH should pose a greater risk necessitating the more secure detention.


Project and Scholarly Work

Summary

In this prospective study, the authors examined whether the DUNDURM – 1, DUNDURM – 3 and DUNDURM – 4 along with other assessment instruments, HCR-20, PANSS, GAF, SRAMM and CANFOR, could distinguish between patients who moved from less secure to more secure in-patient units and vise-versa. Data were gathered as part of a clinical audit of service delivery in a forensic hospital in the Republic of Ireland. A total of 86 male patients were assessed. Patients had a mean age of 40.6 (SD = 12.8) at baseline, and length of stay in the hospital was 7.6 years (SD = 9.9). Primary diagnoses in the sample were schizophrenia (74%), bi-polar affective disorder (10%), schizoaffective disorder (8%), major depressive disorder (3.5%) and intellectual disability (3.5%). A positive move was recorded if there was any move to a less secure unit and a negative move was recorded if there was any move to a more secure unit. Patient movements were documented over a period of 1.07 years. Of the sample, 76.7% had no moves, 12.79% had positive moves and 10.46% had negative moves.

Mean scores on each of the scales were reported for patients admitted into each of the eight different sections of the hospital (e.g., rehabilitation pre-discharge unit, hostel ward, 24 hour nurse care). Total scores on the HCR-20 ranged between 13.3 and 29.0. Patients admitted to the selective adaptive behavior unit (the most secure unit of the hospital) had significantly higher scores on the dynamic (total C and R composite, M = 14.0) and total scales of the HCR-20 (M = 29.0). However H items did not significantly vary as a function of location (range 10.7 to 15.0). While patients who had positive moves had significantly higher scores on the H items of the HCR-20 compared to those with negative or no moves (15.2 vs. 14.3 and 12.4), dynamic items did not significantly differ across groups (6.9 vs. 9.7 and 6.3) although patients with negative moves had higher scores on dynamic items.

Binary logistic regression indicated that for positive moves the location at baseline, DUNDURM-1, HCR-20 dynamic and PANSS general symptoms scores were associated with subsequent positive moves. The receiver operating characteristic was significant for the DUNDURM-1 while ANOVA co-varying for both location at baseline and HCR-20 dynamic score was significant for DUNDURM-1. For negative moves, a binary logistic aggression indicated that location at baseline, DUNDURM-1 and HCR-20 dynamic scores were associated with subsequent negative moves, along with DUNDURM-3 and PANSS negative symptoms in some models. The receiver operating characteristic was significant for the DUNDURM-4 recovery and HCR-20 dynamic scores with DUNDURM-1, DUNDURM-3, PANSS general and GAF marginal. ANOVA co-varying for both location at baseline and HCR-20 dynamic scores showed only DUNDURM-1 and PANSS negative symptoms were associated with subsequent negative moves. The authors concluded that overall the DRUMDRUM-1 triage security score and the HCR-20 dynamic risk measures were best associated with subsequent positive and negative moves. These findings were consistent with past literature that has shown the HCR-20 is useful in predicting a negative move from the community back into a secure hospital.


Summary

This study investigated whether the care pathway was supported by housing patients with the highest levels of risk and psychopathology in high secure units and the lowest levels of risk and psychopathology in the low secure units. Many hospitals and forensic units operate under a care pathway model with several levels of security that patients are progressed through from high to medium to low secure units. The current study was conducted at a large hospital that housed two high secure units, two medium secure units, a low secure unit, and an open unit. Data were collected in a one month period in which 75 men were housed on the units, yet only 70 patients were available for the study. Participants were rated on the dynamic items (C and R scales) of the HCR-20, the Positive and Negative Syndrome Scale (PANSS), the Global Assessment of Functioning (GAF), the Camberwell Assessment of Need Forensic Version (CANFOR), and the Health of the Nation Scales – Secure (HoNOS-Secure).

The HCR-20 C scale (F = 7.7, p < .001), R scale (F = 5.8, p < .001) and total dynamic items score (F = 9.2, p < .001) all stratified significantly across the units. Other measures that also stratified across the units were the HoNOS-Secure, the PANSS totals score, the PANSS positive, and the GAF. The PANSS negative and the CANFOR did not stratify across the levels of security. The authors concluded that the results support the pathways of care model and discussed implications.

Project and Scholarly Work

de Borba Telles, L., Folino, J., & Taborda, J. (2012). Accuracy of the Historical, Clinical and Risk Management Scales (HCR-20) in predicting violence and other offenses in forensic psychiatric patients in
This study examined the predictive validity of the HCR-20 for inpatient aggression in a sample of male forensic psychiatric patients ($N = 68$) in Brazil. Using a concurrent prospective cohort design, the HCR-20 and PCL-R were coded at baseline using information obtained from individual interviews with the patients and their relatives, reviews of therapeutic records, and assessment of prior forensic psychiatry documentation. Episodes of violence and/or antisocial behavior in the hospital were assessed over a one-year period using Yudofsky’s Overt Aggression Scale and Tengstrom’s Follow-Up Questionnaire by a psychiatrist who visited the inpatient unit.

Most participants in the study were Caucasian (73.5%), single (79.4%), with a low level of schooling ($M = 3.74$ years), and employed at the time of their index offense (58.8%). Of the sample, 58.8% had a psychiatric disorder, 20.6% were mentally retarded, 14.7% had a personality disorder, and 5.9% were addicted to alcohol or drugs. During the one-year follow-up period 200 violent or antisocial acts on the ward were recorded, involving 73% of the sample. The mean number of violent/antisocial events per patient was 2.94.

The average HCR-20 total score was 23.32 ($SD = 6.91$), with average scores of 12.35 ($SD = 3.89$), 4.65 ($SD = 2.37$) and 6.32 ($SD = 1.99$) on the H, C, and R subscales, respectively. The average PCL-R total score was 13.54 ($SD = 9.99$), with average scores of 5.78 ($SD = 5.16$) and 6.73 ($SD = 4.26$) on Factor 1 and Factor 2 of the PCL-R, respectively. Predictive power of the HCR-20 was higher for any antisocial event than for a violent offense, whereas the predictive values of the subscales and total score were only moderate. For general offending, AUCLs were 0.82 ($p < .001$), 0.83 ($p < .001$), 0.76 ($p < .01$), and 0.64 ($p > .05$) for total score, H scale, C scale, and R scale, respectively. For violent offending, AUCLs were 0.73 ($p < .05$), 0.65 ($p > .05$), 0.75 ($p < .01$), and 0.69 ($p < .05$) for the total score, H scale, C scale and R scale, respectively. AUCL values were also reported for total, Factor 1 and Factor 2 scores of the PCL-R. For general offending, AUCLs were 0.77 ($p > .05$), 0.69 ($p < .05$), and 0.82 ($p < .001$) for total, Factor 1 and Factor 2 scores, respectively. For violent offending, AUCL values were 0.62, 0.67, and 0.58 (all $p > .05$) for total, Factor 1 or Factor 2 scores, respectively.

Binary logistic regression analyses were performed to determine whether scores on the HCR-20 and PCL-R were significant predictors of any antisocial or any violent event. Scales H, C, R, Factor 1 and Factor 2 were introduced in Block 1 as independent variables, and the totals were introduced in Block 2. With regards to any antisocial event, HCR-20 total scores ($\beta = 1.3$, $p < .001$) and H subscale scores ($\beta = 1.52$, $p < .001$) were found to be predictive. With regards to the perpetration of violence: HCR-20 total scores ($\beta = 1.2$, $p < .05$) and C-subscale scores ($\beta = 1.5$, $p < .05$) were found to be predictive. The same independent and dependent variables were used to calculate the hazard ratios (Cox Regression). Cox regression analyses revealed that scores on the H subscale were the best predictors of any antisocial event, whereas scores on the C subscale were the best predictors of violent events.

The authors concluded that the HCR-20 was suitable for assessing the risk of violence among forensic inpatients in Brazil. This study represents the first evaluation of the predictive accuracy of the HCR-20 in Brazil.

**Project and Scholarly Works**


**Summary**

Aimed to examine the assessment process of experts and staff raters, as well as the predictive ability of these ratings, the present study used an overlapping sample with Dernekiv and colleagues (2001). Eight forensic psychiatric patients (four were also participants in the above mentioned study) were assessed by both expert raters and staff raters using the HCR-20. An average of five staff raters rated each participant. Raters were also asked to indicate the importance of pre-sentence forensic psychiatric reports (a comprehensive report of the patient), case and hospital notes, interview with the patient, and interaction with the patient after making their ratings. Patients were subsequently followed up for 48 months once released into the community.

With regard to the importance of information when scoring the HCR-20, the expert raters considered the pre-sentence forensic psychiatric reports and the interview with the patient more important than the staff raters, who considered interaction with the patient as more important.

No significant differences were seen between the total score or subscale scores between expert and staff raters. For expert raters, the mean scores were 26.38 ($SD = 4.44$), 14.13 ($SD = 3.83$), 6.75 ($SD = 1.58$), and 5.63 ($SD = 1.51$) for the total score, H scale, C scale, and R scale, respectively. For staff raters, the mean scores were 25.88 ($SD = 4.06$), 15.25 ($SD = 3.99$), 5.75 ($SD = 1.58$), and 4.88 ($SD = 1.55$) for the total score, H scale, C scale, and R scale, respectively.
Predictive validity was then assessed for each rater group. For inpatient violence and staff raters, the total score ($r = .63$) and R scale ($r = .76$) were significant predictors. For inpatient violence and expert raters, the total score ($r = .66$) and H scale ($r = .73$) were significant predictors. For violent recidivism and staff raters, none of the scores were significant predictors. For violent recidivism and expert raters, the total score ($r = .81$) and H scale ($r = .81$) were significant predictors. Implications of these results for clinical use are discussed.


**Summary**

This was a reliability study of the HCR-20. Six clinicians each rated six patients on the HCR-20. Reliability coefficients ranged from .76 to .96.


**Summary**

This report discussed the findings of Dernevik (1998), Dernevik et al. (2002), Dernevik et al. (2001), and Dernevik (2004a) that are summarized separately in the annotated bibliography. Broader implications from this group of studies are discussed. (This report also includes another study that is not included, or relevant, to the annotated bibliography.)


**Summary**

The main goal of this study was to evaluate issues related to the process of risk assessment as it pertains to the HCR-20. Specifically, the study evaluated whether “expert” HCR-20 raters (psychologists) differed in their scores from psychiatric nurses. Second, analyses were conducted to determine the extent to which HCR-20 ratings were influenced by clinicians’ feelings towards the patient. The contextual grounding for this approach was drawn from the larger clinical and social psychological literature on biases and heuristics in decision-making.

A total of 8 male patients and 40 clinicians (psychiatric nurses) took part. On average, each patient was rated by five clinicians, and each clinician rated one patient. These patients had serious violent index offences (homicide, rape, assault, arson) and severe mental disorders, as well as personality disorders. They were on average 28 years of age.

The “Feeling Word Checklist” (FWC) was used for clinicians to rate their reactions to the patients they assessed. The FWC is based on a circumplex model with 30 items comprising four dimensions and eight scales, as follows: (1) Helpfulness vs. Unhelpfulness; (2) Closeness vs. Distance; (3) Accepting vs. Rejecting; and (4) Autonomous vs. Rejecting. The FWC predicted HCR-20 scores with Mult. $R^2 = .66$, with feeling Close and Accepting relating to higher scores, and Helpfulness and Autonomy relating to lower scores.

The mean score for the nurse was 26.3 ($SD = 6.1$), whereas it was lower for the “expert” raters ($M = 22.7; SD = 6.5$).

As Dernevik et al. point out, the question of whether the relationship between feelings and HCR-20 scores is evidence for biases in clinical decision-making is not clear. There were no outcome data (i.e., subsequent violence). Further, it is possible that clinicians’ feelings are correlative rather than causative of the HCR-20 ratings. Dernevik et al.’s findings, however, emphasize the importance of limiting biases and over-emphasis on personality to the greatest extent possible, and also the potential importance in professional training on the outcome of an assessment. Further, item bias was not directly assessed (i.e., differential item functioning using Item Response Theory).

**See Also**


**Project and Scholarly Work**

Risk assessment and different risk management levels using the HCR-20. Psychology, Crime, and Law, 8, 83-111.

Summary

This was a prospective study of short-term inpatient and community violence. Participants were 54 consecutive admissions over two years to a forensic psychiatric unit. Most (n = 48) were male. Mean age was 34.2 (SD = 8.92). Most had violent index offences (assault, n = 16; murder, n = 10; great bodily harm, n = 4; arson, n = 10; sex offences, n = 6; other, n = 8). 29% had an Axis I diagnosis only (mostly schizophrenia); 14% had Axis II only; 27% had both; 9% had other combinations of diagnoses.

Predictive analyses were carried out for the whole sample, as well as across three risk management levels: Level one: (High RM) Time spent on a high security ward with no access to the community. Level two: (Medium RM) Time spent living in the hospital but with limited access to occupational and recreational activities in the community. Level three: (Low RM) Time spent in less secure living arrangements and having access to the community while still being monitored regularly.

For overall analyses, HCR-20 effects with inpatient violence were as follows: HCR-20 Total Score (r = .32; AUC = .68); H Scale (r = .37; AUC = .68); C, R, and PCL:SV did not predict inpatient violence. For community violence re-conviction analyses, HCR-20 Total Score AUC = .84; PCL:SV AUC = .71. The C Scale had the highest AUC of the subscales, at .79.

In the low, medium, and high risk management conditions, the measures were most predictive in low and medium conditions, and less to in the high risk management condition. In the High Risk Management condition, only the H Scale was predictive (AUC = .67). HCR-20 Total Score predicted with r = .21 and AUC = .64. C, R, and PCL:SV did not predict. In the Medium Risk Management condition, effects were as follows: HCR-20 Total Score (r = .41; AUC = .82); H Scale (r = .34; AUC = .83); C Scale (r = .36; AUC = .75). R and PCL:SV were not significantly associated with violence, though had small/moderate effect sizes. In the Low Risk Management condition, HCR-20 Total Score (r = .50; AUC = .71); H Scale (r = .48; AUC = .75); R (r = .49; AUC = .62); C and PCL:SV did not predict.

Dernevik et al. interpreted their results as supporting the predictive validity of the HCR-20 for inpatient and community violence. The finding that the HCR-20 was less strongly related to violence in the High Risk Management than in the Medium or Low Risk Management categories, or in the community follow-up, was interpreted not as lack of validity but as effective intensive clinical risk management in this category. This is consistent with the finding and conclusion reached by Müller-Isberner et al. (1999).

Project and Scholarly Works


Summary

This study examined the reliability and validity of the START for predicting inpatient aggression in a sample of 120 male patients in secure psychiatric hospitals in Western Canada. The sample used in this study was primarily Caucasian (75.8%), diagnosed with schizophrenia spectrum disorders (85%), had a comorbid substance use disorder (52.5%), and were in the hospital as result of being found NCRMD (89%). The START, HCR-20, and PCL: SV were coded by graduate research assistants blind to outcome data which had been collected in a previous study through retrospective file review (see Nicholls et al., 2009). Outcome data were coded from hospital files over a 12-month period using the Overt Aggression Scale and were separated into verbal aggression, physical aggression towards objects and physical aggression towards others. About half of the sample (54.2%) engaged in aggression during the follow-up period. The most common form of aggressive behavior was verbal (52.5%), followed by physical aggression towards others (22.5%) and physical aggression against objects (16.7%). Despite high base rates of aggression, aggressive behaviors were generally mild to moderate in terms of severity. Data reported in subsequent analyses reflect dichotomous coding of the presence or absence of the different forms of aggression.

Mean scores were 18.46 (SD = 8.08) and 16.82 (SD = 8.07) on the START strength and vulnerability scales, respectively. HCR-20 mean scores were 13.82 (SD = 3.41), 4.81 (SD = 2.51), 6.19 (SD = 2.36) and 24.90 (SD = 6.59) on the historical, clinical, risk management and total scales. Mean PCL-R scores was 11.72 (SD = 4.20). Inter-rater reliability for the risk instruments used in the study was good. ICC values calculated on a subset of 24 cases were .77 .70, .93, .95, and .85 for PCL: SV, HCR-20, START strength, vulnerability and total scores respectively. ICC value for the Overt Aggression Scale was calculated on a subset of 40 patients and was also high in the study, ICC’s for each of the subscales ranged from .67 to .84.
Convergent and divergent validity of each of the measures was examined. Correlations were moderate to strong between START vulnerability and strength total scores and HCR-20 subscale and total scores. With regards to the START vulnerability total scores, correlations were 0.46, 0.81, 0.77 and 0.83 between the historical, clinical, risk management and total scores of the HCR-20 (all $p < .01$). With regards to the START strength total scores, correlations were -0.43, -0.72, -0.76 and -0.77 between the historical, clinical, risk management and total scores of the HCR-20 (all $p < .01$). PCL-SV was only significantly correlated with START vulnerability total score ($r = 0.21, p < .05$) and the historical ($r = 0.34, p < .01$) and total scores of the HCR-20 ($r = 0.25, p < .01$). There was a significant, large negative correlation between strength and vulnerability total scores on the START ($r = -0.87, p < .01$).

AUC values for the START violence risk estimates were not significantly different than, though somewhat larger than, the HCR-20 violence risk estimates and the PCL: SV total scores for any aggression (AUC = 0.80 vs. AUC = 0.79 and AUC = 0.75), verbal aggression (AUC = 0.78 vs. 0.74 and 0.74), physical aggression against objects (AUC = 0.84 vs. AUC = 0.70 and AUC = 0.63), and physical aggression towards others (AUC = 0.85 vs. AUC = 0.77 and AUC = 0.74), respectively. Cohen’s Kappa on the categorical final risk judgments of the START and HCR-20 was .77. There were no instances were one patient was identified as high risk on one instrument and low risk on the other, and vice versa.

The authors conducted two sets of hierarchical logistic regression analyses to determine whether START assessments added incremental validity over historical risk factors. First, whether START strength and vulnerability total scores and final risk estimates added to the capacity of the historical subscale scores of the HCR-20 to predict aggression were examined. Across models, predictive capability improved significantly, however the models differed regarding whether the strength or vulnerability total scores added incremental validity. For any aggression and verbal aggression, vulnerability total scores added incremental predictive utility, whereas for physical aggression towards others, strength total scores demonstrated predictive utility. Neither vulnerability nor strength total scores added unique contributions to the prediction of physical aggression against objects, though the overall model was significant. For all four outcomes, the addition of START violence risk estimates produced increases in predictive capacity; however, after entering strength and vulnerability scores historical risk factors lost their contribution in the model. Second, whether START strength and vulnerability total scores added to the capacity of the PCL: SV to predict aggression was examined. The same pattern of results was obtained as that found in the previous set of analyses with the HCR-20. The authors concluded that START assessments performed as well as, and sometimes better than assessments using the HCR-20 and PCL: SV, although they did not compare the START to the total score, C or R scales of the HCR-20.


Summary

Due to the limited research specifically focused on dynamic risk factors and lack of research using short follow-up intervals, this study was undertaken to examine the predictive validity of the START and HCR-20 using a repeated assessment design. Participants were 30 male forensic psychiatric inpatients that had been found Not Criminally Responsible on account of Mental Disorder. The participants were selected from a larger sample and classified as aggressive ($n = 15$) and non-aggressive ($n = 15$). The two groups were matched on the type of index offence, length of stay, and age. Participants were assessed using the START and HCR-20 every three months for a year, and outcome data was gathered for each of the assessment periods using the Overt Aggression Scale (OAS).

The mean HCR-20 score was 24.97 ($SD = 6.58$). On the START the mean scores were 18.30 ($SD = 7.52$) for the Strength scale, and 18.94 ($SD = 8.52$) for the Vulnerability Scale. Significant group differences were found between the two groups in terms of their HCR-20 total scores ($t = -8.79, p < .001$). Group differences were also seen on the Strength ($t = 8.70, p = <.001$), and Vulnerability ($t = -8.45, p < .001$) scales of the START.

Overall both of the instruments were found to be predictive of violence. AUCs for the HCR-20 ranged from .72 to .86, depending on the assessment and follow-up period. AUCs for the START ranged from .67 to .87. Survival analyses further verified that the HCR-20 and START were predictive of violence. Moreover, the C and R scales added incremental validity to the H scale. The results are discussed in terms of the implications for the inclusion of dynamic risk factors and the use of instruments to assess the dynamic nature of risk.

See Also

Desmarais, S. L., Wilson, C. M., Nicholls, T. L., & Brink, J. (2010, March). Reliability and validity of the Short-
clinical psychologist or psychotherapists. The group leaders were a diverse group with most having relevant higher vocational or academic training.

The mean HCR-20 scores were: Total score = 26.1 (SD = 6.5), H-scale = 14.6 (SD = 3.3), C-scale = 5.3 (SD = 2.2), R-scale = 6.1 (SD = 2.1). The inter-rater agreement was measured by ICC. The ICC between all three groups for the HCR-20 Total score was .79. For the H-scale the ICC was .82, the C-scale was .64, the R-scale was .57 and the final clinical risk judgment ICC was .65. Inter-rater agreements between subgroups of raters were equivalent to that of all three groups together.

In terms of differing scores by rater type, Group leaders rated significantly lower scores on the H-scale, Risk management items, and HCR-20 Total scores. There were no significant differences in the mean scores between the researchers and the treatment supervisors except for structured clinical risk judgments. Treatment supervisors more often judged patients as “low risk” compared to researchers.

Researchers stated that they spent about 120 minutes per risk assessment, group leaders spent about 30 minutes and supervisors about 15 minutes per assessment. Also, researchers stated that they based their assessments predominantly on file information, whereas group leaders and treatment supervisors mostly relied on personal experiences with the patient.

Correlations between HCR-20 scores and a measure of feelings towards the patients showed many significant correlations. The HCR-20 total score was correlated with measures of patient’s: helpfulness ($r = -.28; \ p < .01$), unhelpfulness ($r = .38; \ p < .01$), distant ($r = .2; \ p < .05$), accepting ($r = -.19; \ p < .01$), rejecting ($r = .34; \ p < .01$), and controlled ($r = .46; \ p < .01$). The HCR-20 risk judgment was correlated with measures of patient’s: helpfulness ($r = -.34; \ p < .01$), unhelpfulness ($r = .33; \ p < .01$), close ($r = .19; \ p < .5$), distant ($r = .4; \ p < .01$), accepting ($r = -.23; \ p < .01$), rejecting ($r = .34; \ p < .01$), and controlled ($r = .37; \ p < .01$).

Stepwise multiple regression analyses showed that feelings of being controlled or manipulated by a patient significantly predicted high HCR-20 scores. 21% of the variance in the HCR-20 Total score was explained by feelings of being controlled by the patient. Also in stepwise regressions, feelings that the patient was close and distant predicted high risk judgments; in contrast, feelings that the patient was helpful predicted low risk judgments. Together these three explained 23% of the variance in risk judgments.

**Project and Scholarly Work**


**Summary**

This research project assessed whether clinicians and researchers differ in their violence risk assessment of the same patients and whether raters’ feelings towards the patients play into their risk assessments. This study used the Dutch version of the HCR-20 on 60 patients (53 men and 7 women) in a Dutch psychiatric forensic hospital. The groups which coded the HCR-20 were comprised of 5 independent researchers, 7 treatment supervisor and 32 group leaders. The treatment supervisors were mostly

Summary

This prospective study examined differences in accuracy between researchers (n = 9), treatment supervisors (n = 8), and group leaders (n = 59) with respect to individual versus consensus ratings and structured final risk versus actuarially based risk judgments. The sample comprised 127 men (a subset of whom were reported on previously; see de Vogel & de Ruiter, 2004) whose mean age at admission was 32.9 (SD = 9.6, range = 17-66). The index offenses were: 44% (attempted) homicide, 33% sexual offenses, 16% other violent offenses such as Robbery, 7% arson. Mean length of stay in the hospital was 3.7 years (SD = 2.4, range = 0-12). More than half of the participants had abused substances in the past (8% alcohol, 15% drugs, and 44% multiple substances) and most had received previous psychiatric treatment.

Participants varied in terms of their treatment phase at the time the HCR-20 was coded. For participants who were commencing their first unsupervised leave from the hospital (n = 9), entering the transmural treatment phase (n = 28), or already were in the transmural treatment phase (n = 24), the R-scale was coded for the outside context. For participants who were newly admitted to the hospital (n = 49) and for existing inpatients (n = 17), the R-scale was coded for the context inside (risk of inpatient violence).

Raters coded the HCR-20 between January 2001 and June 2004 for each case independently and agreed upon a consensus score and a final risk judgment during a case conference. For 19 (15%) patients, more than one HCR-20 rating was completed because there was a change in their treatment phase. The most recent risk assessment was used for those participants.

Outcome data were obtained from daily information bulletins published in the hospital that report on inpatient violence and violence that occurred outside the hospital (e.g., for patients who were in the transmural treatment phase). The definition of physical violence was the same as that used in the HCR-20 manual. The mean follow up period was 21.5 months (SD = 10.9, range = 1-37). For individuals under mandated treatment conditions, data on violent recidivism was not obtained after the court order expired (n = 20; mean follow up period after discharge for this subgroup = 15 months, SD = 8.8, range = 4-34).

Group leaders gave significantly lower total and R-scale scores (p < .05) compared to researchers and treatment supervisors. There were no significant differences in mean HCR-20 scores between researchers and treatment supervisors. The mean HCR-20 consensus scores were higher (but not significantly so) than the mean HCR-20 scores of the three individual rater groups. Mean total scores were: researchers = 26.1 (SD = 6.1), treatment supervisors = 25.8 (SD = 6.1), group leaders = 24.1 (SD = 5.8), consensus = 26.8 (SD = 5.6). Mean H-scale scores were: researchers = 14.5 (SD = 3.1), treatment supervisors = 14.3 (SD = 3.4), group leaders = 14.0 (SD = 3.4), consensus = 14.8 (SD = 3.1). Mean C-scale scores were: researchers = 5.3 (SD = 2.1), treatment supervisors = 5.3 (SD = 2.2), group leaders = 5.0 (SD = 2.0), consensus = 5.5 (SD = 2.1). Mean R-scale scores were: researchers = 6.3 (SD = 2.2), treatment supervisors = 6.2 (SD = 2.2), group leaders = 5.3 (SD = 2.2), consensus = 6.4 (SD = 1.9).

There were no significant differences between the rater groups in final risk judgments. The percentages of low HCR-20 final risk judgments were: 24% researchers, 30% treatment supervisors, 21% group leaders, and 28% consensus. The percentages for judgments of moderate risk were: 45% researchers, 46% treatment supervisors, 43% group leaders, and 48% consensus. The percentages for judgments of high risk were: 31% researchers, 24% treatment supervisors, 35% group leaders, and 24% consensus.

AUC values for physical violence for the total score were: researchers = .79 (SD = .05), treatment supervisors = .81 (SD = .05), group leaders = .75 (SD = .05), consensus = .85 (SD = .04). AUC values for the H-scale were: researchers = .73 (SD = .06), treatment supervisors = .74 (SD = .06), group leaders = .75 (SD = .06), consensus = .77 (SD = .05). AUC values for the C-scale were: researchers = .76 (SD = .06), treatment supervisors = .75 (SD = .05), group leaders = .66 (SD = .06), consensus = .80 (SD = .05). AUC values for the R-scale scores were: researchers = .74 (SD = .06), treatment supervisors = .71 (SD = .05), group leaders = .63 (SD = .07), consensus = .79 (SD = .05).

AUC values for the final risk judgment were: researchers = .77 (SD = 2.2), treatment supervisors = .75 (SD = .05), group leaders = .64 (SD = .07), consensus = .86 (SD = .04). Group leaders compared to researchers had a significantly lower AUC value for the final risk judgment (χ² (1, N = 127) = 6.3, p < .01). Group leaders’ ratings compared to consensus ratings were significantly lower for the C-scale, R-scale, total score, and final risk judgment (χ² (1, N = 127) = respectively 6.8, 4.9, 4.6 and 20.1, p < .05). The AUC value for the HCR-20 consensus final risk judgment was significantly higher than the individual final risk judgment of researchers, treatment supervisors and group leaders (χ² (1, N = 127) = respectively 6.9, 5.3, and 20.1, p < .01).
Correlations for the HCR-20 total score were: researchers = .35, treatment supervisors = .36, group leaders = .30, consensus = .43. Correlations for the H-scale were: researchers = .27, treatment supervisors = .28, group leaders = .29, consensus = .32. Correlations for the C-scale were: researchers = .31, treatment supervisors = .31, group leaders = .19, consensus = .36. Correlations for the R-scale were: researchers = .29, treatment supervisors = .27, group leaders = .16, consensus = .35. Correlations for the final risk judgment were: researchers = .35, treatment supervisors = .33, group leaders = .19, consensus = .49. All p values < .01 for consensus, researchers, and treatment supervisors and at least < .05 for group leaders (except R-scale, p = .16).

Participants who scored above the median (27) relative to those below the median had significantly more had significantly more incidents of physical violence (Kaplan Meier log rank = 15.8, p < .001; odds ratio = 21.6, 95% CI = 2.8-167.2). Cox regression analyses with the three scales entered on the first block and final risk judgment entered on the second using the forward conditional method resulted in a significant model fit ($\chi^2(3, N = 127) = 22.9, p < .001$) at Block 1. HCR-20 final risk judgment demonstrated incremental validity as there was significant improvement to the model’s fit at Block 2 ($\chi^2$ change (1, N = 127) = 6.8, $p < .01$).

AUC values and Pearson correlations were used to examine the predictive validity of consensus ratings for physical violence of the HCR-20 items. Items 2, 4, 5, and 7 from the H-scale, items 11, 12, 14, and 15 from the C-scale, and items 16, 17, and 19 from the R-scale had significant AUC values and correlations. Significant AUC values ranged from .67-.74 and significant correlations ranged from .21-.32. Cox regression analysis with all items included yielded a significant model fit ($\chi^2(20, N = 127) = 43.7, p < .01$). Using the forward conditional method to determine which HCR-20 items were significant predictors of incidents of physical violence produced a final model in which items 2 ($e^B = 6.4, 95\% CI = 1.5-28.0$), 15 ($e^B = 3.4, 95\% CI = 1.5-8.1$), and 17 ($e^B = 3.4, 95\% CI = 1.2-10.0$) were significant predictors of incidents of physical violence.

The HCR-20 total score and final risk judgments were significantly predictive for both verbal abuse (total score: AUC = .72, SE = .05, $r = .36, p < .01$; final risk judgment: AUC = .65, SE = .05, $r = .28, p < .01$) and verbal threat (total score: AUC = .79, SE = .05, $r = .36, p < .01$; final risk judgment: AUC = .71, SE = .05, $r = .31, p < .01$).


Summary

This research project assessed the reliability and predictive validity of the HCR-20 and the SVR-20 (Sexual Violence Risk-20). The project also assessed who would be the most suitable to perform risk assessments. The study used 60 patients (53 males, 7 females), assessed them initially before their entrance into a transmural phase and then again in the transmural phase.

Inpatients’ H, C, R and Total scores were higher than those in the transmural phase. Inpatient final risk judgments were higher than when in the transmural phase.

Project and Scholarly Works


Summary

In this paper, the authors present gender-sensitive risk assessment guidelines for female (forensic) psychiatric patients, the Female Additional Manual (FAM). The FAM is a manual designed to be used alongside the HCR-20 and was developed on the basis of a literature review and clinical expertise. Additional considerations for rating ten of the original HCR-20 items for women are provided, and nine specific risk factors for women were added, such as Prostitution, Parenting difficulties, Pregnancy at young age and Covert behavior. Furthermore, two new coding aspects are included in the FAM: marking critical items; and judging the risk of self-harm, victimization and non-violent
offending in addition to risk of violence to others.

In 2010, a prospective study was carried out on the psychometric properties of the FAM in a Dutch forensic psychiatric hospital admitting both men and women. The FAM, in addition to the HCR-20, HCR: V3, PCL-R and the SAPROF, was coded prospectively for 42 women and a matched group of 42 men. Information related to incidents during treatment, such as violence towards other and self, victimization and criminal offending were also recorded.

Interrater reliability coded on a subset of 20 cases indicated good interrater reliability for total score and individual items of the FAM. IRR values were .95 for FAM total score, and ranged between .63 and 1.00 for all new items. For final risk ratings interrater reliability was moderate to good. IRR values were .95, .85, .54 and .73 for violence towards others, self-destructive behavior, victimization and non-violent criminal behavior, respectively.

Codings on the FAM and other instruments for the female sample were compared to those of the male sample in order to assess the specific applicability of the FAM items for women. Women had higher scores on 7 of the 9 FAM items and men had higher scores on 2 of the 9 FAM items (psychopathy and problematic behavior during childhood). Overall, women scored higher on psychiatric factors and men scored higher on antisocial factors. With regards to violence against others, AUC values were .76, .66, .83, .71, and .87 for FAM total, H, C, R, and final risk ratings, respectively. With regards to self-destructive behavior, AUC values were .80, .68, .77, .81 and .97 for total, H, C, R, and final risk ratings, respectively. AUC values for the final risk judgments were .63 and .99 for victimization, and non-violent criminal offending, respectively. AUCs were not reported separately for the HCR-20 and the FAM. The authors concluded that the FAM was promising for assessing not only violence to others, but self-destructive behavior in a female sample and may be a useful addition to the HCR-20.


### Summary

The predictive validity of the Dutch version of the HCR-20 was examined in a forensic psychiatric sample of 42 women admitted between 1985 and 2003. A sample of 42 male forensic psychiatric patients, also admitted between 1985 and 2003 and matched on birth year, type of index offense, ethnicity, and type of psychopathology, was used as a comparison group.

The HCR-20 was coded on the basis of file information. For the women, ratings were made retrospectively for 15 cases and prospectively for 27 cases.  Good interrater reliability was observed for the women for the total score, H-scale, and final risk judgment (n = 27; ICCs = .75, .82, .74) and moderate for the C-scale and R-scale (ICCs = .55, .51). For the men, half the ratings were retrospective and half were prospective. Good interrater reliability was observed for the total score, H-scale, C-scale, and final risk judgment (n = 28; ICCs = .77, .82, .70, .69).

There were significant mean differences between the genders on several HCR-20 items but the total and scale scores were comparable. For women, mean scores were: HCR-20 total (25.9, SD = 5.5); H-scale (14.0, SD = 2.9); C-scale (5.4, SD = 2.0); R-scale (6.6, SD = 1.9). For men, mean scores were: HCR-20 total (27.1, SD = 6.5); H-scale (14.9, SD = 3.0); C-scale (5.4, SD = 2.3); R-scale (6.8, SD = 2.1). With respect to the HCR-20 final risk judgments, women were judged as moderate risk significantly more often, whereas men were judged as high risk significantly more often. The three most frequently coded ‘other considerations’ differed for each gender. For men they were financial problems, lack of prospects for the future, and violent fantasies whereas for women they were forming a new intimate relationship, care for children, and prostitution. Analyses of the predictive validity included two types of violence collapsed into a single outcome variable: (1) violent recidivism (operationalized with the HCR-20 definition of violence) after discharge was obtained from official judicial records for the “retrospective participants” and (2) data on inpatient violence was obtained from daily hospital information bulletins that detailed any disruptive incidents (incidents were coded only if they were acts of physical violence directed towards other persons). Values for all HCR-20 indices were higher for men than women. For men, AUC values for HCR-20 total and scale scores ranged from .75 to .88 and rs ranged from .42 to .62. For women, AUCs ranged from .52 to .63 and rs ranged from .07 to .22. Values for final risk judgments were lower for women than for the HCR-20 total and scale scores across both men (AUC = .91, r = .70) and women (AUC = .86, r = .57). Predictive indices for the PCL-R generally were lower than for the HCR-20.

Results indicate the predictive ability of the HCR-20 may be maximized when judgments of final risk are used rather than an actuarial approach wherein individual risk factors are summed.


Abridged abstract (English translation of the study not available):

Background: Violence perpetrated by women is a growing problem. Research has shown that the risk factors associated with women differ from those associated with men and that the risk assessments currently in use are not adequate for predicting violence in women. *Aim:* To develop a clinically relevant, useful tool for an accurate, gender-sensitive assessment of risk of violent behaviour in women and to offer guidelines for risk management in women. *Method:* On the basis of literature research, clinical expertise and the results of a pilot study, we adapted the much-used ‘Historical Clinical Risk management-20’ (HCR-20) for use with female (forensic) psychiatric patients who have a record of violence towards other people. *Results:* The ‘Female Additional Manual’ (FAM) supplemented and added value to the HCR-20 for assessing the risk of violent behaviour by women. *Conclusion:* The FAM is a valuable addition to the currently available risk assessment tools in that it provides a more accurate gender-specific risk assessment with regard to female (forensic) psychiatric patients. Future research will have to further demonstrate the value of the FAM.

Project and Scholarly Works


Summary

This prospective study examined how changes in the dynamic risk factors of the HCR-20 and dynamic protective factors of the SAPROF influenced violent outcome over the course of clinical treatment. The study sample consisted of 325 high risk violent and sexually violent offenders in a forensic psychiatric hospital in the Netherlands. Scores on the C and R items of the HCR-20 and dynamic items of the SAPROF were analyzed at different stages during treatment (intramural, supervised leave, unsupervised leave, transmural, discharge). Violent incidents during treatment were recorded over a 12 month period.

Results show that as patients moved through different treatment stages, their risk level as reflected by their combined risk and protective factor scores, reduced accordingly. The combined effect of decrease in risk factors and increase in protective factors was a strong predictor for the decline in incidents of violence over the course of clinical treatment. AUCs were .77, .79, .81, and .76 for the Total SAPROF, HCR-20 and combined HCR-SAPROF scores and final risk judgments, respectively. Overall, the study results indicate the usefulness of dynamic risk and protective factors for informing effective clinical treatment and for measuring changes in individual risk levels over time.


Summary

The authors investigated the predictive validity, inter-rater reliability and survival rates while using the HCR-20 and PCL-R. The sample consisted of 120 patients discharged from a Dutch forensic psychiatric hospital between 1993 and 1999. The patients had a mean duration of treatment of 58.7 months and there was an average follow-up period of 73 months for this study. There were four different ways of discharge for these patients: transmural (N = 30; termination of treatment by court in line with hospital’s advice and after a resocialization phase), conform advice (N = 30; termination of treatment by court in line with hospital’s advice without resocialization phase), contrary to advice (N = 30; termination of treatment by court against the hospital’s advice) and readmission to another institution (N = 30; readmission to another institution).

Inter-rater reliability was measured using ICC’s. The ICC for the HCR-20 total score was .83. For the H-scale it was .89, for the C-scale it was .76, for the R-scale it was .58, and for the structured final risk judgment it was .73.

The mean scores for the HCR-20 and PCL-R by type of discharge were as follows. Transmural means were: PCL-R total (15.4), HCR-20 total (22.8), H-scale (12.6), C-scale (3.7), R-scale (6.5). Conform means were: PCL-R total (17), HCR-20 total (22.8), H-scale (12.8), C-scale (4.3), R-scale (5.6). Contrary means were: PCL-R total (20.2), HCR-20 total (27.6), H-scale (14.6), C-scale (5.4), R-scale (7.6). Readmission means were: PCL-R total (25.3), HCR-20 total (32), H-scale (16), C-scale (7), R-scale (9.1).

For the H-scale there were significant differences between the transmural and conform means as compared to the contrary mean (p < .05) and the contrary mean as compared to the readmission mean (p < .05). For the C-scale there
were significant differences between the transmural and conform means as compared to the contrary and readmission means ($p < .05$). For the R-scale there were significant differences between the transmural and conform means as compared to the contrary and readmission means ($p < .05$). For the HCR-20 total score there were significant differences between the transmural and conform means as compared to the contrary and readmission means ($p < .05$). For the PCL-R total score, there were significant differences between the transmural and conform means as compared to the contrary mean ($p < .05$) and the contrary mean as compared to the readmission mean ($p < .05$).

Significant differences were found in the level of risk judgments given across the four discharge types. For the HCR-20, low risk judgments were given significantly more often to transmural and conform groups than to the readmission group ($p < .05$). Use of the HCR-20 also led to more moderate risk judgments for the transmural, conform and contrary groups as compared to the readmission group ($p < .05$). Lastly for the HCR-20, this measure led to more high risk judgments for the transmural and conform groups as compared to the contrary and readmission groups ($p < .05$). Using a cut-off of 26 on the PCL-R, there were higher judgments of risk given to those in the contrary and readmission groups as compared to the transmural or conform groups ($p < .05$).

Results showed that there were no significant differences between the transmural and conform or contrary groups in terms of violent recidivism. The conform group had a lower recidivism rate for violent offenses ($p < .05$), and the readmission group had a higher recidivism rate for violent offenses than the other three groups ($p < .01$).

The predictive validity of the HCR-20, PCL-R and clinical judgment for violent offending were calculated using AUCs and Pearson’s correlations. AUC’s: HCR-20 total score (.82; $p < .001$), H-scale (.80; $p < .001$), C-scale (.77; $p < .001$), R-scale (.79; $p < .001$), Risk judgment (.79; $p < .001$), PCL-R total score (.75; $p < .001$), PCL-R with cut-off of 26 or greater (.65; $p < .01$) and unstructured clinical judgment (.68; $p < .01$). Correlations: HCR-20 total score (.52; $p < .01$), H-scale (.47; $p < .01$), C-scale (.46; $p < .01$), R-scale (.47; $p < .01$), Risk judgment (.51; $p < .01$), PCL-R total score (.43; $p < .01$), PCL-R with cut-off of 26 or greater (.39; $p < .01$) and unstructured clinical judgment (.32; $p < .01$).

The authors conclude that the HCR-20 structured final judgment was significantly more accurate than unstructured clinical judgment in predicting violent recidivism ($p < .05$). The HCR-20 was also significantly more accurate than the PCL-R in predicting violent recidivism ($p < .05$) except when the item H7 (psychopathy) was removed from the HCR-20 total score ($p = .08$).

Summary

This retrospective study provides the first validation of the Structured Assessment of Protective Factors for violence risk (SAPROF), an instrument developed as a strengths-based supplement to the HCR-20 (and other risk assessment tools), in a sample of 126 male violent offenders discharged from intensive forensic psychiatric hospital treatment in the Netherlands. The study authors hypothesized that the combined use of both the HCR-20 and the SAPROF would increase predictive validity over either instrument alone. An overall total score of risk and protection was composed by subtracting SAPROF total score from HCR-20 total score, resulting in a total risk score corrected for available protective factors (HCR-SAPROF total score). The dependent variable, recidivism, was defined as any new conviction for a violent offense according to the HCR-20 definition of violence (actual, attempted, or threatened violence). For all patients, follow-up time started on the day of discharge and was recidivism was recorded at 1 year intervals over a 3 year period. Twenty patients were readmitted and their recidivism data could not be retrieved, thus they were excluded from subsequent predictive validity analyses. Criminal records showed that eight of the 105 discharged patients were reconvicted of a violent offense within 1 year, 15 patients after 2 years, and 20 patients within 3 years.

The sample used in this study consisted of violent offenders discharged between 1990 and 2006 from a TBS hospital. The average treatment length was 5.3 years ($SD = 2.2$) and the mean age at release was 31 ($SD = 7.3$). The majority (83%) of the sample had an Axis II personality disorder. A history of substance abuse was present in 65% of the cases. HCR-20 and SAPROF were coded from patient’s files, which contained biographical information, psychological reports, reports to the court regarding treatment progress, treatment plans and treatment evaluations, by trained researchers, including the three study authors. All raters were blind to previous risk assessments and recidivism outcome data. Based on a subsample of 40 cases ICC values were 0.88 and 0.85 for total SAPROF scores and final protection judgments respectively. All individual factors had moderate to excellent inter-rater reliabilities (range .42 to .94). Mean scores were: SAPROF Total $M = 11.65$ ($SD = 6.41$); Internal factors $M = 3.48$ ($SD = 1.84$); Motivational Factors $M = 5.01$ ($SD = 3.91$); and External Factors $M = 3.18$ ($SD =
1.89). The Final Protection Judgment was low for 41% of the sample, moderate for 51% of the sample and high for 8% of the sample. Means scores on the HCR-20 total and subscales, as well as SPJ ratings were not reported. There was a significant negative correlation between the HCR-20 and the SAPROF ($r = -.69$, $p < .01$). The relationship between the HCR-20 and the SAPROF total scores was calculated for the entire sample ($N = 126$). Analyses showed a high negative correlation between both instruments ($r = -.69$, $p < .01$). The highest inter-item correlations were found between the SAPROF factor Self-control and the HCR-20 factors Impulsivity ($r = -.73$, $p < .01$) and Unresponsiveness to treatment ($r = -.69$, $p < .01$) and between the SAPROF factor Motivation for treatment and the HCR-20 factor Noncompliance with remediation attempts ($r = -.67$, $p < .01$).

Receiver operating characteristic analyses were conducted for the 1, 2, and 3 year follow-up periods. SAPROF total scores showed good predictive validity for non-recidivism of a violent offenses, AUC values were large at 1-year follow-up (AUC = .85) and 2-year follow-up (AUC = .80) and moderate to large for 3-year follow-up (AUC = .74). For all three follow-up periods, the predictive validity for the violent recidivism of the HCR total score was lower than the predictive validity for non-recidivism of the SAPROF total score. However, this difference was not significant. AUCs were .81, .77, and .68 at 1-, 2-, and 3-year follow-up, respectively. With regards to the combined HCR-SAPROF total score AUC values were .85, .81, and .72 for 1, 2, and 3 year follow-up respectively. Comparative analyses on the AUC values showed a significantly better predictive validity of the HCR-SAPROF measure over the HCR-20 total score for both 1-year and 3-year follow. The SAPROF Final Protection Judgment (FPJ) and the Integrated Final Risk Judgment (IFRJ) showed good predictive validity for (no) violent recidivism with 1 and 2 years follow-up (FPJ AUC = .82 and .77, respectively, IFRJ AUC = .80 and .72, respectively). However, predictive validities of both final judgments decreased at 3-years after discharge (FPJ AUC = .71; IFRJ AUC = .65). The best individual predicting SAPROF factors for no violent reconvictions were Self-control (AUCs = .83, .74 and .73 at years 1, 2, and 3, respectively) and Work (AUCs = .83, .76 and .71, respectively, both $p < .01$).

The analyses of the correlations between the total score on each instrument and recidivism ($n = 105$) revealed significant negative results for 1-, 2-, and 3-year follow-ups for the SAPROF ($r_{pb} = -.35, -.38$, and $-.35$, respectively, all $p < .01$) and significant positive results for all three follow-up times for the HCR-20 ($r_{pb} = .31, .34$, and .25, respectively, all $p < .05$).

The results demonstrate that combined total score the HCR-SAPROF was a significantly better predictor of violent reconvictions than the HCR-20 total score, at least at the 1- and 3-year follow-ups. These findings suggest that future risk assessments would benefit from combination of risk and protective measures for predicting future violence.


**Summary**

The current study investigated the value of assessing risk and protective factors when determining level of risk in a sample of discharged forensic psychiatric patients. In addition, the study examined whether protective factors were similar for both violent offenders and sexual offenders. The HCR-20 and SAPROF were coded retrospectively for a sample of 188 patients with a history of violent ($n = 105$) or sexual offending ($n = 83$) who had been admitted to a Dutch forensic psychiatric hospital under a TBS order. The study sample was predominately Caucasian and had either an Axis II personality disorder (66%) or traits (20%). A history of substance use problems was present in 72% of cases.

HCR-20 and SAPROF were coded from patient files by nine trained raters blind to recidivism outcome. Interrater reliability was calculated on a subsample of 24 cases. ICC values were .74 for the HCR-20 total score and .79 for the SAPROF total score. In addition to calculating total scores on each scale, the authors composed an overall total score of risk and protection by subtracting the SAPROF total score from the HCR-20 total score (HCR-SAPROF). Recidivism was the dependent variable in the study. Recidivism was operationalized as any new conviction after discharge for a violent (sexual or non-sexual) offense according to the HCR-20 definition of violence (any actual, attempted or threatened violence). All patients were followed up with in the community for a period of at least three years after discharge ($M = 11.1$). To compare predictive validities at fixed follow-up times, official reconvictions within one and three years after release were used. Violent recidivism rates were 8% for one year, 19% for three years, and 30% for long-term follow-up for the violent offender sample. For the sexual offender sample, this was 7, 17 and 45%, respectively.

The authors tested whether there was a moderating effect of offense type (violent vs. sexual) on the relationship between total scores on the HCR-20, SAPROF, and the HCR-20 SAPROF and re-offending. The interaction between offense type and total scores were entered into a logistic regression analysis for each tool, with new convictions for violent
offending at the different follow-up times as outcome. In all cases, offense type did not significantly moderate the relationship between total score and violent recidivism indicating that risk and protective factors operated in similar ways for both offender types. Samples were pooled together for further analysis.

Analysis of the relationship between post-treatment HCR-20 total score and SAPROF total score showed a negative correlation between both instruments ($\rho = -.76, p < .001$). The analyses of the correlations between the total scores on the tools and recidivism revealed significant negative results for one year, three year and long-term follow-up for the SAPROF ($r_{pb} = -.32, -.35$, and -.39, respectively, all $p < .001$) and significant positive results for all three follow-up times for the HCR-20 ($r_{pb} = .33, .32$, and .26, respectively, all $p < .001$) and for the combined index of HCR-SAPROF ($r_{pb} = .34, .35$, and .34, respectively, all $p < .001$). When controlling for the HCR-20 in a partial correlation analysis, the correlation between the SAPROF and recidivism remained significant at both three year ($r_{pb} = -.18, p < .05$) and long term ($r_{pb} = -.31, p < .001$) follow-up. On the other hand, when controlling for the SAPROF correlations between the HCR-20 and recidivism were no longer significant (values not reported by the authors).

ROC curve analyses were used as an index of predictive validity. The SAPROF total score showed good predictive validity (AUC one year = .85; AUC three year = .75; AUC long-term = .73). The dynamic protective factors were the strongest predictors of desistance from violence, even at long-term follow-up (AUC one year = .86; AUC three year = .75; AUC long term = .72). The predictive validity of the HCR-20 total score for violent recidivism was comparable to that of the SAPROF total score for one and three-year follow-up (AUC = .84 and .73, respectively). However, the long-term predictive accuracy of the risk factors (AUC = .64) was not as strong as for the protective factors. The dynamic (Clinical and Risk management) risk factors predicted future violence better than the static (Historical) ones. The combined total score of the HCR-SAPROF index was the best predictor of violent recoventions for one and three-year follow-up (AUC = .87, .76, respectively), although not significantly better than either the SAPROF or HCR-20 alone. Comparative analyses on the AUC values showed that at long-term follow-up the HCR-SAPROF index total score (AUC one year = .87; AUC three year = .76; AUC long term = .70) predicted violent recidivism significantly better than the HCR-20 total score alone ($\chi^2(1, N= 188) = 13.4, p < .001$), however, at one year and three-year follow-up these differences were not significant.

To further assess incremental predictive validity of the SAPROF protective factors over the HCR-20 risk factors, and the interaction between risk and protective factors over the independent total scores on both tools, the authors conducted hierarchical logistic regression analysis on each of the three follow-up times. Although not significant for one-year follow-up, for both three-year and long-term follow-up statistically significant improvements to the model were found when the SAPROF was added. In addition, partial correlation analysis showed that both three year and long term follow-up the correlation between violent outcome and protective factors remained significant after controlling for risk scores. The relationship between Final Protection and Final Risk judgments and recidivism was also examined. In general, patients with the highest risk levels recidivated the most. Overall, especially within the moderate risk and high-risk group patients with higher levels of protection at discharged showed less violent recidivism. Taken together, results demonstrate good predictive validity of dynamic risk and protective factors over time and the incremental validity protective factors in assessing (desistance from) violence risk.

**Project and Scholarly Works**


**Abridged abstract (English translation of the study not available):**

In the present study, 64 sex offenders in Switzerland were retrospectively rated with the PCL-SV, the HCR-20 +3 and the SVR-20. These participants are part of a larger study by the Forensic Department of the Psychiatric University. The risk assessments were coded based on prior risk assessment reports as well as criminal reports. The scores on the PCL-SV, HCR-20 +3 and the SVR-20 were compared to prior scores on the Structured Risk Assessment of Basel. Results of this study confirm the utility of PCL-SV, HCR-20+3 and SVR-20 in a German-speaking sample of sex offenders. The authors conclude the risk assessment instruments should be used primarily with antisocial and physically aggressive sex offenders.

**Project and Scholarly Works**

**Dolan, M., & Blattner, R. (2010).** The utility of the Historical Clinical Risk -20 Scale as a predictor of outcomes in decisions to transfer patients from high...
to lower levels of security-A UK perspective. BMC Psychiatry (Open Access Journal).

Summary

This pseudo-prospective study examined the predictive validity of the HCR-20 for outcome following transfer in a sample of high security forensic patients discharged via medium security care in the UK ($n = 72$). The mean age of the sample was 36.4 years ($SD = 11.5$). A majority of the sample was male (87%), Caucasian (79%), met the criteria for substance abuse dependence (61%) and diagnosed with schizophrenia (67%). Primary index offenses leading to admission were violence against others (64%), sexual offenses (17%), and arson or criminal damage (19%). Outcomes following transfer were classified as either “successes” or “failures.” Failure was based on any of the following criteria: direct return to the high security unit, return to the high security unit after discharge to the community or reconviction for a serious offense (i.e., murder, manslaughter, assault, rape, indecent assault, Robbery, or arson) after discharge to the community. Overall, 55.5% patients had an outcome that was classified as a failure based on the assigned categories. Of these patients, 46% returned directly to the high-security hospital. Of the patients discharged to the community (46%), 21% were reconvicted, 15% for violence against persons.

The HCR-20 was scored from comprehensive case file information at the time of discharge. Item H7 (PCL-R score) was not coded for this study. Three types of follow-up outcome data were collected blind to the initial HCR-20 scores: (1) reconvictions were coded from the Home Office Offender Index; (2) readmissions to district and forensic hospitals (readmissions could be of several types, including those under the Mental Health Act that reflected concern over an escalation in violence secondary to a relapse in mental state); and (3) self/collateral reports of violence were coded from community mental health teams’ computerized records. Violence included sexual violence, punching, biting, choking, kicking, or assault with a weapon that resulted in physical injury to the victim.

Mean HCR-20 scores, with the psychopathy item omitted, were: total = 19.37 ($SD = 5.7$, range = 9-31); H-scale = 11.82 ($SD = 3.65$, range = 6-18); C-scale = 3.34 ($SD = 2.20$, range = 0-10); R-scale = 4.07 ($SD = 1.45$, range = 1-7).

The mean length of stay in the community was 59 months ($SD = 37.64$). There was a significant negative relationship between time in the community and HCR-20 total score ($r = -.48, p < .0001$), H-scale ($r = -.29, p < .014$), C-scale ($r = -.57, p < .001$), and R-scale ($r = -.37, p < .001$).

For the following analyses, median splits were performed and high and low scores refer to scores above and below the median, respectively. No significant associations were detected between high and low total scores on the HCR-20 and re-offending ($\chi^2 = 2.71, p = .10$) or violent re-offending ($\chi^2 = 1.72, p = .18$). Likewise, none of the subscales demonstrated a significant association with reconviction. However, the number of readmissions was correlated significantly with HCR-20 total score ($r = .40, p < .0001$), C-scale ($r = .26, p < .026$), and R-scale ($r = .31, p < .007$), but not H-scale. Chi-square analyses indicated a significant association between high total scores and all subscales for any form of readmission and especially for readmission under the Mental Health Act (MHA). For incidents of violence reported by participants or their carers, there was a significant association with high HCR-20 total scores ($\chi^2 = 10.19, p < .006$), H-scale ($\chi^2 = 16.13, p = .001$), and C-scale ($\chi^2 = 7.46, p < .008$). This association nearly reached significance for the R-scale ($\chi^2 = 3.06, p = .08$).


Summary

This study investigated the predictive validity of the HCR-20 total and subscale scores among violent patients ($N = 70$) discharged to the community under fairly intensive supervision. Participants were discharged between 1992 and 2000 and had stayed in the hospital for a mean of 24 months ($SD = 14.49$). Most of the sample was Caucasian (83%) and single (80%). The mean age at admission was 35.3 years ($SD = 10.12$). The most common Axis I diagnosis was schizophrenia (73%) and roughly one-third of the sample had either primary or secondary diagnoses of a PD (the most common being APD, 26%). Almost half (44%) had a history of co-morbid substance misuse.

The HCR-20 was scored from comprehensive case file information at the time of discharge. Item H7 (PCL-R score) was not coded for this study. Three types of follow-up outcome data were collected blind to the initial HCR-20 scores: (1) reconvictions were coded from the Home Office Offender Index; (2) readmissions to district and forensic hospitals (readmissions could be of several types, including those under the Mental Health Act that reflected concern over an escalation in violence secondary to a relapse in mental state); and (3) self/collateral reports of violence were coded from community mental health teams’ computerized records. Violence included sexual violence, punching, biting, choking, kicking, or assault with a weapon that resulted in physical injury to the victim.

Mean HCR-20 scores, with the psychopathy item omitted, were: total = 19.37 ($SD = 5.7$, range = 9-31); H-scale = 11.82 ($SD = 3.65$, range = 6-18); C-scale = 3.34 ($SD = 2.20$, range = 0-10); R-scale = 4.07 ($SD = 1.45$, range = 1-7).

The mean length of stay in the community was 59 months ($SD = 37.64$). There was a significant negative relationship between time in the community and HCR-20 total score ($r = -.48, p < .0001$), H-scale ($r = -.29, p < .014$), C-scale ($r = -.57, p < .001$), and R-scale ($r = -.37, p < .001$).

For the following analyses, median splits were performed and high and low scores refer to scores above and below the median, respectively. No significant associations were detected between high and low total scores on the HCR-20 and re-offending ($\chi^2 = 2.71, p = .10$) or violent re-offending ($\chi^2 = 1.72, p = .18$). Likewise, none of the subscales demonstrated a significant association with reconviction. However, the number of readmissions was correlated significantly with HCR-20 total score ($r = .40, p < .0001$), C-scale ($r = .26, p < .026$), and R-scale ($r = .31, p < .007$), but not H-scale. Chi-square analyses indicated a significant association between high total scores and all subscales for any form of readmission and especially for readmission under the Mental Health Act (MHA). For incidents of violence reported by participants or their carers, there was a significant association with high HCR-20 total scores ($\chi^2 = 10.19, p < .006$), H-scale ($\chi^2 = 16.13, p = .001$), and C-scale ($\chi^2 = 7.46, p < .008$). This association nearly reached significance for the R-scale ($\chi^2 = 3.06, p = .08$).
ROC curve analyses were used as another index of predictive accuracy that the HCR-20 total score for readmission under the MHA was .85 (p < .001, CI = .76 -.95). AUC values also were significant for self/collateral reports of violence (AUC = .76, p < .001, CI = .65-.87) and re-offending (AUC = .71, p < .05, CI = .56-.87). The AUC value for serious re-offending was not significant (AUC = .67, p = .15, CI = .47-.88). All three scales predicted readmission under the MHA (AUC values ranging from .65 to .78), with highest values obtained for the H-scale.

Kaplan-Meir survival analyses revealed significant relationships between above-median HCR-20 scores and poor survival in the community. Log rank values were: MHA readmission = 27.73 (p < .001); self/collateral reported violence = 17.14 (p < .001); re-offending = 3.85 (p < .05); and violent re-offending = 5.08 (p < .05). The authors noted (but did not provide a quantitative summary) that when survival analyses were used to examine time at risk in the community, they found evidence that the C-scale and R-scale outperformed the H-scale.

The discussion section highlights reasons that may have contributed to the finding of a significant relationship between HCR-20 scores and readmission but not reconviction (e.g., high level of supervision, hospital policy pertaining to re-admittance at time of deterioration in mental state/increase in risk of violence).

**Project and Scholarly Work**


**Summary**

In this study, the authors investigated the psychometric properties of the Part A baseline assessment component of the Violence Risk Scale second edition (VRS-2; Wong & Gordon, 2000). Participants were 136 male inpatients at the Edenfield Medium Secure Unit in Manchester, UK who were admitted to the unit between 1995 and July 2003. The VRS was coded based on admission notes. The HCR-20 was completed on a subsample of 80 cases that were then followed-up in 12 months.

The mean age of the sample was 35.5 years (SD = 9.45). The majority were Caucasian (80.1%). Primary diagnoses included schizophrenia (76.4%), schizoaffective disorder (10.3%), affective disorder (3.7%), personality disorder (4.4%), and organic brain syndrome (0.7%). The majority of participants had been referred from prison (55.1%). However, 18.4% had been transferred from a high security hospital, 14.7% from non-forensic district services, 10.3% from the courts or probation services, and 1.5% from the forensic community service. The vast majority (97%) had been charged with or convicted of a criminal offence. Of these, 34.6% had an index offence of wounding, 14.0% homicide, 12.5% arson, 7.4% sexual assault, 8.1% theft, 2.2% public order offences, and 2.9% possession of offensive weapons. A further 15.4% had multiple index offences recorded, all of which included a violent offence. Four had no criminal conviction but had been admitted because of unmanageable violence in district services.

Demographic information and some historical information was collected from case files. These same files, admission summaries, index forensic assessment reports, and pre-admission court reports (available at admission) were used to score the VRS-2 and the HCR-20. The VRS-2 and the HCR-20 were scored independently by researchers, blind to each other’s ratings. Data on outcome (episode of physical violence towards others) were examined by a third researcher to avoid any potential bias. The VRS-2 contains an item assessing institutional violence: this item was rated based on violence in other settings prior to the index admission to the secure unit. The inter-rater reliability of the VRS-2 was based on a comparison of 23 cases rated previously by another rater and the current rater. The intraclass correlation coefficients were satisfactory, with alphas of .96, .85, and .89 for the VRS-2 static, dynamic, and total scores respectively.

The VRS-2 total scores had a mean of 41.0 (SD = 11.3). The static item scores had a mean of 8.17 (SD = 3.8). The dynamic item had a mean of 32.9 (SD = 8.9). The means of the HCR-20 were: Total M = 20.5, SD = 6.1, 4-32; H subscale M = 10.1, SD = 3.5, 2-18; C subscale M = 5.91, SD = 1.88, 1-10; R subscale M = 4.52, SD = 1.58, 0-8. Correlations between the two measures were all highly significant. Participants who had engaged in institutional violence during the 12 month follow-up period had higher mean VRS-2 total, subscale scores, HCR-20 and subscale scores than the non-violent group. Comparison of the predictive accuracy of both measures indicated that they had moderate predictive accuracy (VRS-2 AUCs = .62-.72; HCR-20 Total, H, C, R AUCs = .72, .66, .73, .67). Overall, the dynamic scales in both measures had greater predictive accuracy that the more static scales. A logistic regression analysis indicated that the subscale scores from both measures were significant contributors to the prediction of institutional violence, however, only the C subscale was a significant predictor in the forward entry model.
Summary

The predictive validity of the HCR-20 (version 2) was evaluated prospectively among a complete caseload of patients \((N = 47)\) managed by a community forensic team. The sample primarily was male \((n = 43; 91\%)\) and African-Caribbean \((n = 35; 74\%)\). Many participants had a history of violence in the community \((n = 43; 91\%)\) or in an inpatient setting \((n = 23; 49\%)\).

Data for all participants were collected over a three-month period by the author, who had worked clinically with some of the patients, via file review and an interview with each participant’s key worker (who typically was a community psychiatric nurse). No direct contact with the participant occurred in the course of data collection. The PCL-R was completed for a subset of participants \((n = 33)\). Mean HCR-20 scores were: Total \((M = 21.65, SD = 6.15)\); Historical \((M = 13.40, SD = 3.31)\); Clinical \((M = 4.11, SD = 2.32)\); Risk Management \((M = 4.33, SD = 2.27)\). The range of mean item scores on the Historical scale was 1.19 \((H10)\) to 1.87 \((H6)\). The range of mean item scores on the Clinical scale was .54 \((C3)\) to 1.22 \((C1)\). The range of mean item scores on the Risk Management scale was .41 \((R1)\) to 1.48 \((R2)\).

Recidivism data were collected 2.5 years after the HCR-20s were scored. Outcome data were based on file records and information collected from clinical staff. Eight participants were charged or convicted of a new offence. Mean total scores of recidivists \((M = 29.4)\) and non-recidivists \((M = 21.2)\) were statistically significant \((p < .05, \text{independent } t\text{-test})\). Re-offending of two participants appeared to be linked closely to deteriorated mental state. Among the other six participants, all of whom maintained their mental stability, re-offending appeared to be related to instrumental violence, substance misuse, and antisocial personality characteristics. Implications for targeting specific types of patients for forensic services (versus generic services) were discussed.

Project and Scholarly Work


Summary

Because of a purported lack of research investigating the use of risk assessment tools with offenders with intellectual disabilities \((ID)\), the authors investigated the use of the VRAG and the HCR-20 in medium secure unit with a sample of 25 participants with a diagnosis of a learning disability \((n = 25)\) and a control group of mentally disordered offenders \((n = 45)\). Participants were subsequently followed for 6 months, with 80% of the learning disabled group committing at least one violent incident and 40% of the control group committing at least one violent incident.

For the learning disabled group the mean score on the VRAG was 14.60 \((SD = 7.23)\), and mean scores on the HCR-20 were 26.60 \((SD = 4.54)\), 14.92 \((SD = 2.18)\), 7.28 \((SD = 1.99)\), and 4.36 \((SD = 1.80)\), for the total scores, H, C, and R scale respectively. For the control group the mean score on the VRAG was 6.38 \((SD = 10.90)\), and mean scores on the HCR-20 were 23.71 \((SD = 5.98)\), 14.29 \((SD = 3.55)\), 5.60 \((SD = 2.45)\), and 3.89 \((SD = 2.19)\), for the total scores, H, C, and R scale respectively.

The VRAG was found to predict institutional violence better in the learning disabled group \((AUC = .87)\) than the control group \((AUC = .60)\). The HCR-20 also yielded larger effects in the learning disabled group with AUCs of .77, .77, .66, .73, and .88 for the total scores, H, C, R scale, and final risk judgments. In comparison, the AUCs were .58, .42, .67, .62, and .63 for the control group.

The main conclusion drawn from the present study was that both the VRAG and HCR-20 are excellent predictors of future violence in offenders with intellectual disabilities.

Project and Scholarly Works


Summary

Predicting institutional violence in offenders with intellectual disabilities. In W. R. Lindsay (Chair), Risk assessment for offenders with intellectual disability \((ID)\), Symposium conducted at the annual conference of the International Association of Forensic Mental Health Services, Edinburgh, Scotland.

Project and Scholarly Work

Fitzgerald, S., Gray, N. S., & Snowden, R. J. (2009, June). Predicting institutional violence in offenders with intellectual disabilities. In W. R. Lindsay (Chair), Risk assessment for offenders with intellectual disability \((ID)\), Symposium conducted at the annual conference of the International Association of Forensic Mental Health Services, Edinburgh, Scotland.
The accuracy of actuarial predictions of inpatient violence using a cut-off score of 27 on the HCR-20 versus clinicians’ structured professional judgments (SPJ) was compared. The HCR-20 was administered by three doctoral level psychologists within the first week of arrival to 169 patients (138 men and 31 women) admitted consecutively to a state hospital between February 2002 and January 2003. The most common admission diagnoses were schizoaffective (18%) and paranoid schizophrenia (16%). The inter-rater reliability coefficient for 12 cases was .94. Episodes of inpatient violence (operationalized by the definition of violence in the HCR-20 manual) were recorded from hospital event records for a minimum of three months post-admission.

For actuarially derived predictions, the hit rate = 71%, sensitivity = 30% (12/40), specificity = 86% (95/111), positive predictive power (PPP) = 43% (12/28), negative predictive power (NPP) = 77% (95/123), and AUC = .61 (range: .51-.72). SPJ-based predictions (patients were rated either as high risk or low/moderate risk) generally were higher: hit rate = 77%, sensitivity = 45% (18/40), specificity = 88% (105/120), PPP = 55% (18/33), NPP = 83% (105/127), and AUC = .70 (range: .56-.77).

A step-wise regression was completed using the number of violent inpatient episodes as the criterion variable and overall HCR-20 scores and five-level SPJ predictions (low, low-moderate, moderate, moderate-high, and high) as the predictor variables. SPJ-based predictions added incremental validity over actuarial predictions (an increase in $r^2$ from .036 to .092, $p < .05$), whereas the reverse was not true.

The study also reports on clinicians’ predictions regarding the situational contexts in which violence might occur for each participant based on his or her historical background. Results provide support for the use of the SPJ approach in making predictions of inpatient violence among forensic psychiatric patients.


Summary

The authors noted that the majority of studies on the HCR-20 have used samples that comprise predominantly Caucasians of European heritage. The purpose of the present study was to examine retrospectively cultural differences in violence risk assessment of psychiatric inpatients using the HCR-20. Participants were drawn from a sample of 169 consecutive admissions (the same sample reported on by Fujii, Lichton, & Tokioka, under review). Participants were included in this study if they described themselves as Asian-American (AA; $n = 51$), Euro-American (EA; $n = 46$), or Native American of part-Hawaiian (NAH, $n = 38$) heritage. Participants were considered AA if their ethnicity was Japanese, Chinese, Korean, Filipino, or Vietnamese and NAH if they reported Hawaiian as one of their ethnic languages. Participants with a mixed ethnic heritage, apart from the NAH group, were excluded. The final sample consisted of 88 men and 20 women and had a mean age of 40.1 years ($SD = 12.6$) and a mean education level of 11.9 years ($SD = 2.5$).

There were no differences in rates of institutional violence (i.e., threats or assaults on patients and staff) among the three ethnic groups. ROC analyses indicated the highest accuracy for predicting inpatient violence was obtained for the NAH group ($AUC = .730$) and the lowest accuracy for the AA group ($AUC = .575$; $AUC$ for the EA group = .638). Stepwise multiple regressions were conducted for each ethnic group using HCR-20 items as predictor variables and the number of violent events (multiplied by log10 to control for a skewed distribution) as the criterion variable. Results indicated a unique pattern of predictors was associated with each cultural group. Models for AA and EA each produced a single significant predictor. For AA, item C4 (impulsivity) accounted for 16.1% of the variance. For EA, item H2 (young age at first violence) accounted for 13.3% of the variance. The largest effect size ($R^2 = .430$) was obtained for NAH, which had three significant predictors (H2, young age at first violence; H3, relationship instability; and R1, plans lack feasibility).

Results are discussed in terms of possible explanations for the disparities in observed predictive ability of the HCR-20 as a function of ethnicity.

Project and Scholarly Work


Summary

The authors compared the clinical, criminal, and personality characteristics of male forensic psychiatric patients with schizophrenia who were categorized into high and low psychopathy groups on the basis of PCL:SV scores. The HCR-20 was coded, but it was not a main focus of this study. Participants ($N = 61$) were recruited from two
forensic psychiatric hospitals in England. Their mean age was 37.79 years (SD = 8.52); data on the racial composition of the sample was not provided.

Criminal history information (i.e., number of offences, type of offences, and age at first offence) was coded dichotomously from official conviction records reported within case files. Data were collected on the following measures: PCL:SV, PANSS, HCR-20, a self-report measure entitled the Antisocial Personality Questionnaire (APQ; Blackburn & Fawcett, 1999), and the Chart of Interpersonal Reactions in Closed Living Environments (CIRCLE; Blackburn & Renwick, 1996). The authors reported that “trained researchers completed the psychopathy, risk and symptom related assessments based on file review and interview where appropriate.” It was stated that the PCL:SV was completed using file review and interview, but the basis of HCR-20 scores was not specified. Ratings on the CIRCLE were made by nursing staff.

A researcher blind to baseline assessment data recorded episodes of institutional aggression using computerized official incident reports covering the period from admission to assessment. An episode was defined as aggressive “if the patient was the clear instigator or co-aggressor, and if the incident involved verbal or physical aggression to the staff, patients or property.” Length of follow-up was not specified.

The mean PCL:SV score was 12.5 (SD = 5.37). Patients were classified as psychopathic if they scored above the 75th percentile on the PCL:SV (total score of 16 or higher). The psychopathic patients (n = 19) had a higher mean total HCR-20 score (M = 25.61, SD = 5.38) than the non-psychopathic patients (n = 42) (M = 19.29, SD = 5.49), (t (57) = -4.09, p < 0.001). The psychopathic patients also had significantly higher mean scores on the Historical scale (M = 15.7, SD = 1.87), (t (57) = -5.09, p < 0.001). To avoid criterion contamination, data were analysed with and without the HCR-20 items H7 Psychopathy and H9 Personality Disorder. Using the total and Historical scores adjusted on this basis, the psychopathic patients group still had higher total (t (57) = -3.05, p < 0.01), and Historical (t (57) = -2.52, p < 0.05) scores. The psychopathic group also had significantly higher Clinical (M = 5.44, SD = 2.48) scores than the non-psychopathic group (M = 3.90, SD = 2.69), [t (57) = -2.07, p < 0.05]. No differences between the psychopathic (M = 4.44, SD = 2.91) and non-psychopathic (M = 3.49, SD = 2.20) groups were observed on Risk Management scores, (t (57) = -1.39, n.s.).

Predictive validity data were provided for the PCL:SV, but not the HCR-20. The psychopathic group was more likely to have engaged in an episode of institutional aggression (n = 13, 72%) than the non-psychopathic group (n = 13, 34.2%; χ² = 7.1, p < 0.01) between admission and time of assessment. The psychopathic group also had a significantly higher mean number of aggressive incidents in the first year of admission (M = 2.74, SD = 4.92; mean rank = 36.5) than the non-psychopathic group (M = 0.18, SD = 2.18; mean rank, 28.51), Mann Whitney U = 294.5, p < 0.05. There was a significant difference between the psychopathic (M = 17.17 months, SD = 31.68) and non-psychopathic (M = 28.48 months, SD = 42.24) survival curves for time in months following admission to first aggressive incident (Kaplan Meier, Log rank statistic [df 1] = 7.64, p < 0.01).

The authors concluded that their findings generally were consistent with previous research that has examined associations between psychopathy scores and violence risk and criminality in general, as well as in patients with schizophrenia. They also noted that assessing personality functioning, including interpersonal style, may help in developing appropriate treatment interventions to mitigate the impact of such personality pathology on maladaptive behaviours such as poor compliance and institutional aggression.

---

**Project and Scholarly Works**


**Summary**

Using a pseudo-prospective case note analysis, this study examined whether the HCR-20 was equally effective for the prediction of future violence across varying mental health diagnoses in a sample of mentally disordered offenders discharged from medium secure facilities in the UK. The total sample consisted of 996 male patients with a mean age at discharge of 37.7 years (SD = 9.2). Primary diagnoses were schizophrenia (63.8%), mood disorder (11.2%), substance misuse (13%), personality disorder (18.0%), mental retardation (12.9%) and other diagnoses (8.4%). A majority of the sample were White (69.2%). The mean length of stay within the medium secure service was 436 days (SD = 510).

Psychiatric diagnoses were determined using ICD-10 criteria by a consultant psychiatrist upon patient admission to the hospital. Two psychologists completed the HCR-20 using file-based information which included full criminal record history, admission and discharge reports, social work and probation information, and nursing records. All HCR-
20 assessments were completed blind to study outcome. Inter-rater reliability was 0.80, 0.92, 0.90 and 0.85 for the total, historical, clinical and risk management scales of the HCR-20, respectively. The dependent variable in the study was the occurrence of an offense, grouped as either any or violent, following discharge from the hospital. Convictions were obtained from the Home Office Offenders’ Index.

Overall, the authors were able to score the HCR-20 and obtain follow-up information for 890 of the patients. Of the sample, 19.4% were convicted of an offense within 2 years of discharge from the unit and 11.3% were convicted of a violent offense. Looking at the sample as a whole, the HCR-20 total score, historical scale and the risk scale produced significant predictions, but the clinical scale did not. AUCs were 0.73, 0.72, 0.55 and 0.70 for the total, historical, clinical and risk management scales of the HCR-20. With respect to any conviction, the AUC values were 0.69, 0.69, 0.51 and 0.68. Again, HCR-20 total, historical and risk scales produced significant predictions, but the clinical scale was not significant.

Patients were stratified according to whether they had received a particular mental health diagnosis. Because there was not sufficient data for some subgroups to perform meaningful analysis, only groups with a sample size larger than 100 (i.e., schizophrenia, personality disorder, substance use, mental retardation and mood disorder) were examined in subsequent analyses. The authors found that the HCR-20 was a significant predictor of future violence in all the psychiatric diagnostic groups they were able to test. However there were some variations in the efficacy of the instrument between groups. HCR-20 total, historical, clinical and risk scores were most predictive in patients with mental retardation (AUCs = 0.80, 0.84, 0.68, and 0.70). Total, clinical and risk scores were the worse for predicting violence among patient with personality disorders (AUCs = 0.62, 0.51 and 0.62). Historical items were the least predictive in patients with a substance use disorder (AUC = 0.60). The same pattern of findings was obtained with respect to general offending. Again, HCR-20 total, historical, clinical and risk scores were most predictive in patients with mental retardation (AUCs = 0.80, 0.79, 0.64, and 0.76). Historical and clinical items were worse for individuals with a substance use disorder (AUCs = 0.62 and 0.45). Total and risk scores were worse for patients with a personality disorder (AUCs = 0.62 and 0.62). The authors concluded that the relatively poor prediction of the HCR-20 for those with personality disorders or substance use disorders might be due to the fact that both of these conditions are associated with impulsive behavior and chaotic lifestyles, thus behavior may be less predictive and more prone to outside influences that are unknown at the time of the evaluation of future risk. However, it should be noted that the HCR-20 tends to be comparably predictive in studies relying on samples consisting primarily of personality disordered people (i.e., offender samples) than it is with forensic or civil psychiatric samples (see meta-analysis by Guy, 2008). As such, these findings might be sample-specific.


Summary

This was a pseudo-prospective study of 887 male forensic psychiatric patients discharged from four medium secure units in the UK between December 1992 and September 2001. Four psychologists completed the HCR-20 based on mental health and criminal justice files and were blind to outcome. The outcome variable was the occurrence of a violent offence or any offence after discharge from the hospital based on information obtained from the UK Home Office. Violence referred to violence against the person including kidnapping, criminal damage endangering life, Robbery, rape and indecent assault. Any offences referred to all offences committed during the follow-up period. Time to offence was calculated as the difference between the discharge date and the time of reconviction for the subsequent offence.

The total sample consisted of 996 male patients with a mean age at discharge of 37.7 years (SD = 9.2, range 16.9-71.2). Of those 996, 887 had an HCR-20 completed. Most patients (69.2%) were White, 21.6% were of Black Caribbean or Black African origin, 2.4% were of Asian origin, 5.2% were of other or mixed ethnicity and 5.2% were of unknown ethnicity. The mean length of stay within the hospital was 436 days (SD = 510 days, range 7-3785 days). The participant’s primary diagnosis was mainly schizophrenia or a psychotic disorder (56.2%) and the rest were mood, personality, mental retardation, developmental or organic disorders with 3.2% being of unknown diagnosis.

Many of the subsequent analyses are based on sub-samples of the overall sample. All sub-samples were compared with the total sample and no significant differences were found in terms of patient characteristics. Inter-rater reliability for the HCR-20 based on 20 cases yielded a collective interclass correlation of .80. The mean scores for the HCR-20 and its subscales are as follows: Total M = 18.3, SD = 6.2; 0-36; H subscale M = 11.3, SD = 3.7, 0-20; C subscale M = 3.2, SD = 2.4; 0-10; R subscale M = 3.7, SD = 2.6; 0-10.

After five years, 34% of participants had a new conviction, with 10% receiving a new conviction for a violent offence. The authors reported AUCs for 6 months to 5 years post-discharge. The HCR-20 was a good predictor of violent offences with AUCs between .70-.76. However, the
predictive accuracy of the HCR-20 (and its subscales) slightly declined over time and this was a statistically significant trend ($p > .05$). The H subscale was also a good predictor (.68–.77) and the R subscale (AUC .63–.69) showed moderate levels of predictive efficacy. In contrast, the C subscale was not predictive of violent offences (AUC .54–.61). The HCR-20 showed similar predictive ability with any convictions, but the AUCs were slightly lower (HCR-20 total, AUC .69–.75; H subscale, .69–.75; C subscale, .51–.55; R subscale, .66–.69). Only the C subscale was not significant.


**Summary**

This study was a pseudo-prospective case-note analysis of the ability of the HCR-20, PCL-R and the VRAG to predict general and violent re-offending in patients with and without intellectual disabilities (ID). The final sample consisted of 1,141 patients released between 1990 and 2001. Participants were admitted to the hospital on the basis of (a) having a serious mental illness, ID, or personality disorder; (b) having been convicted of a criminal offense ($n = 881$); or (c) having exhibited behavior that might have led to a conviction in different circumstances ($n = 260$). The ID group ($n = 145$) all had a diagnosis of mental retardation (MR) and consisted of 121 patients with mild MR, 18 with moderate MR, 5 with severe MR, and 1 with unspecified MR. In the ID group, 49 patients had a diagnosis of ID alone, and 96 patients had a comorbid diagnosis of another mental disorder (either mental illness or personality disorder). The non-ID group ($n = 996$) consisted of all the other participants, all of whom had some form of psychiatric diagnosis but without ID. In the ID group there were 118 (81.4%) men and 27 (18.6%) women, with a mean age at the time of discharge of 31.54 years ($SD = 8.94$, 18.84–65.78). In the non-ID group, there were 843 (85.6%) men and 153 (15.4%) women, with a mean age at the time of discharge of 31.95 years ($SD = 9.28$, 16.90–71.25). The two groups did not significantly differ on gender or age at discharge. The ID group had a lower number of previous convictions ($M = 8.30$, $SD = 13.05$) than the non-ID group ($M = 11.80$, $SD = 16.35$), $t (1139) = 2.47$, $p < .05$.

The scoring of the risk assessments was completed at the point of discharge based on case review notes. Raters were blind to outcome. Not all risk instruments could be completed on all participants because of a lack of relevant file information. Reliability of all instruments was high (VRAG ICC = .95; PCL-SV total ICCs between .89 and .95; HCR-20 total: ICCs between .80 and .88). The ratings were made in a set order of PCL-SV, HCR-20, and then VRAG since the PCL-SV is component of both and to minimize the influence of the more objective VRAG on the more subjective HCR-20.

Outcome information was collected from the Home Office Offenders Index (2000). Violent offenses included all offenses classified as violence against the person by the Home Office, as well as kidnapping, criminal damage endangering life, Robbery, rape, and indecent assault. The ID group had higher VRAG total scores, PCL-SV total scores, Part 1 scores, Part 2 scores, HCR-20 total scores, H subscale scores, and C subscale scores. The ID group was reconvicted at a slower rate (approximately one half) than the rate of the non-ID group for both violent offenses (e.g., after 2 years, 4.8% for the ID group and 11.2% for the non-ID group) and general offenses (e.g., after 2 years, 9.7% for the ID group and 18.7% for the non-ID group). Survival analysis showed these differences to be significant (violence: Log Rank [1] = 7.11, $p < .01$; general Log Rank [1] = 8.19, $p < .01$). The VRAG AUC for predicting violent reconviction after a 5-year follow-up period in the ID group was .74 which was nearly identical to that of the non-ID group.

The PCL-SV was a good predictor of both violent and general reconvictions in the ID group and non-ID group, yielding large effect sizes. The HCR-20 was a very good predictor of violent reconviction in the ID group, achieving an AUC of .79. For general offending, the HCR-20 was again a greater predictor of convictions for the ID group (AUC = .81) than the non-ID group (AUC = .68), and this difference was statistically significant ($p < .05$). The strong performance of the HCR-20 total score was also reflected in the History subscale for the ID group (AUC = .80–.81) but somewhat less so for the Clinical subscale (AUC = .69–.71). The Risk Management subscale did not achieve statistical significance. A series of paired z-score comparisons revealed no significant differences in the predictive accuracy of the VRAG, PCL-SV, or HCR-20 within the ID group or non-ID group.

**Project and Scholarly Work**

Summary

This study compared the predictive accuracy of the HCR-20, PCL:SV, and the Offender Group Reconviction Scale (OGRS; Copas & Marshall, 1998; this is a criminogenic risk assessment tool based on six demographic and offending history variables that estimates the probability of reconviction within 2 years of release) among 315 forensic psychiatric patients discharged from a medium-secure facility in South Wales, United Kingdom between 1992 and 1999. The sample primarily was male (87.6%), Caucasian (84.4%; 12.4% Black Caribbean or Black African; 1.3% Asian; .6% ‘mixed’; 1.3% ‘unknown’), and diagnosed with schizophrenia or psychotic disorder (49.2%; 16.8% personality disorder; 9.8% affective disorder; 6.3% drug induced psychosis; 5.1% MR; 1.0% substance misuse disorder; 3.2% ‘other’; 8.6% ‘unknown’).

Two psychologists blind to outcome completed all assessments using file information available at discharge, which consisted of mental health, criminal, social work, and probation records. Participants were followed up for at least two years (mean = 6.00 years, SD = 1.77 years). During the follow-up period, 36.5% were convicted of any type of offense.

Mean HCR-20 scores were: Total (19.90, SD = 7.02; range 0-36); H-scale (11.39, SD = 3.97; range 0-20); C-scale (3.77, SD = 2.42; range 0-10); R-scale (4.68, SD = 2.63; range 0-10). Mean PCL:SV scores were: Total (8.25, SD = 5.18; range 0-20), Part 1 (3.79, SD = 3.79; range 0-11), and Part 2 (4.50, SD = 2.83; range 0-12). The mean OGRS score was .49 (SD = .29; range .03 -.99). HCR-20 and PCL:SV total and scale/Part scores correlated highly and significantly with one another (ranging from .36 to .78). Correlations for the OGRS with the HCR-20 and PCL:SV tended to be lower and were not consistently significant.

Survival analysis revealed that 87% of the offenses occurred within approximately 3 years. The Mantel-Cox log-rank statistic was used to evaluate the percentage of patients in low, medium, and high risk predictor groups who committed an offense following discharge for the three measures. For the PCL:SV, risk groups were defined as follows: low (scores of 12 or less); medium (scores of 13-17); and high (scores of 18 or more). The distribution of scores was used to trisect the sample into groups for the other two measures. For the HCR-20, groups were defined as follows: low (scores of 16 or less); medium (scores of 17-22); and high (scores of 23 or more). Groups for the OGRS were: low (<.29); medium (.29 -.67); and high (>-.67). Although significant results were obtained with respect to any type of offending outcome for all measures, the log-rank value for the OGRS was much higher (83.78) than the values for the HCR-20 (10.70) and PCL:SV (10.76).

Mean scores on the three measures were compared across participants who offended and those who did not. Cohen’s d values were as follows: HCR-20 total (.35), H scale (.38), C scale (-.08), R scale (.41), PCL:SV total score (.54), Part 1 (.25), Part 2 (.70), and OGRS (1.28).

Using ROC analysis, total scores of all three measures were associated significantly with offending outcome (AUC values for the HCR-20, PCL:SV, and OGRS were .61, .66, and .81, respectively). AUC values for the subscales were more variable, with the H scale (.62), R scale (.62), and PCL:SV Part 2 (.72) reaching significance, but with the C scale (.48) and PCL:SV Part 1 (.57) failing to do so. ROC analyses that examined serious and minor offenses revealed a similar pattern of results. When participants were divided into groups on the basis of diagnosis (i.e., mental illnesses, personality disorders, and ‘other’ diagnoses that included mental retardation, developmental disorder, and physical diagnoses), the size of the AUC values for the mental illness and ‘other’ groups was similar to the above-described values for the overall sample (although none of the values except for the OGRS were significant for the ‘other’ group).

Finally, a logistic regression analysis was undertaken to investigate whether the HCR-20 and/or PCL:SV could make an additional significant contribution to an OGRS-only model. Using a forced-entry method, no total or scale/Part variables added incremental validity.

The discussion section reiterated the findings and noted that the timing at which the C scale was scored (i.e., prior to discharge when symptomatology was as low as it likely ever would be, rather than during a time of active symptomatology) may have impacted the findings. The authors concluded that adoption of a singular focus on mental health factors ignores important sources of information predictive of reoffending.

Project and Scholarly Work


Summary

This study examined the use of the HCR-20 by 10 Australian community forensic mental health services (CFMHS). Each CFMHS location completed a structured questionnaire to obtain comparative data on the use of the HCR-20. During the 12-month survey period, the number of HCR-20 assessments conducted ranged from 6 to 168. Differences in service models impacted on who was seen,
whether reassessments were undertaken, and involvement of generalist mental health staff. Of the 10 locations, 2 assessed only high-risk patients, 6 repeated assessments (either weekly ($n = 1$), every 3 months ($n = 4$) or every 6 months ($n = 1$)), 7 provided preliminary feedback, 6 conducted peer reviews, and 5 discussed assessments with supervisors. All assessments were completed by psychologists, while 70% also involved psychiatrists and nurses, 60% involved social workers and 50% involved registrars. Four of the locations used the PCL-R to code H7 (psychopathy) 100% of the time, while for the other six locations inclusion rates varied from 0 – 90%. Key issues involved in the application of SPJ risk assessments in clinical practice were discussed.

**Project and Scholarly Work**


**Summary**

The HCR-20 is a robust predictor of violence in a variety of settings (e.g., forensic, civil psychiatric). However, there is a lack of research on its effectiveness identifying insanity acquittees who are most at risk of re-arrest/recommitment following their transfer from forensic to civil psychiatric settings. This retrospective study examined the association of HCR-20 risk factors with recommitment of insanity acquittees to forensic hospital. The HCR-20 was coded from patient files of 157 insanity acquittees in eastern US discharged from forensic hospitals to a less restrictive setting. Of these participants, 34.4% were recommitted to the forensic setting between 1977 and 2010. The Historical scale was found to be the best predictor of recommitment over time. In particular, higher scores on factors H2 (young age at first violence), and H10 (prior supervision failure) and lower score on H6 (major mental illness) significantly increased likelihood of recommitment after release from a forensic facility. Only 1 of 10 dynamic items were found to be significant (C2 Negative Attitudes). The authors concluded that the HCR-20 was useful in differentiating those who were recommitted from those who were not.

**See Also**


**Project and Scholarly Works**


**Summary**

This study used a prospective design to determine the relationship between the HCR-20 and levels of security in a forensic psychiatric hospital. Over 12 months, they followed 220 individuals (209 men and 11 women) who had a hospital order sentence and had been hospitalized for the entire 12 month period. They predicted that if risk factors were changed by successful treatment, the dynamic part of the HCR-20 (the CR-10) should decrease.

The following significant correlations between individual C and R scale items and level of security were found over time: C1 ($r = -.286$), C2 ($r = -.264$), C4 ($r = -.236$), C5 ($r = -.347$), R1 ($r = -.42$), R2 ($r = -.443$), R3 ($r = -.237$), R4 ($r = -.409$), R5 ($r = -.227$). The C-scale in total showed a significant correlation with level of security over time ($r = .369; p < .001$), as did the total R-scale ($r = .575; p < .001$) and to a lesser extent, the H-scale ($r = .167; p < .05$). The PCL also showed a significant correlation with level of security over time ($r = .227; p < .01$).

The authors conclude that there are robust correlations between CR-10 items and levels of security and that the CR-10 seems to be a good indicator of treatment progress.

**Scholarly Works**


**Summary**
Using the same data as reported above (Gretenkord et al., 2002), the authors sought to clarify the relationship between the dynamic items on the HCR-20 and treatment success. The mean C scale score was 5.3 (SD = 2.2) and the mean R scale score was 7.6 (SD = 1.9). A consistent pattern was seen across the 8 levels of security and the mean C and R scale scores. That is, as the level of security increased, so too did the mean C and mean R scale score. The C scale and R scale scores were correlated with level of security ($r = .37$ and $r = .52$, respectively). A similar pattern was seen at the item level with almost all of the C and R items correlating with level of security, as reported above. The HCR-20 total score was also correlated with level of security ($r = .40$). The authors concluded that all the C and R items and total scores can be used to predict level of security. Moreover, they discussed that the findings provide some reassurance to the decision making of the clinicians in charge of security levels, in that a reduction in security level is associated with a lower score on the C and R scales.


Summary

100 forensic psychiatric patients were rated on the German version of the HCR-20 (which includes 3 variables not in the original version). There were 96 men, and the mean age of the sample was 38.8 years. Only the H and C scales were rated. Most index offences were of a violent nature: homicide (24%); severe bodily harm (21%); violent sexual offences (20%); arson (13%); and 24 other offences. Close to half (43%) of the sample had primary diagnoses of functional psychosis.

Two psychiatrists rated a subsample of 45 offenders, allowing interrater reliability analyses. For the H Scale items, Kappa ranged from .54 to 1.00, with a mean Kappa of .89. In 91% of cases, the two clinicians were within one point on ratings of H Scale total scores. Kappa was not as high for the C Scale, ranging from .33 to .65, with a mean Kappa of .49. In 71% of cases, clinicians were within one point on the C Scale.

Mean H scores were greatest for personality disordered patients with low IQs ($M = 13.6$) and lowest for patients with major brain damage ($M = 9.5$). Homicide offenders ($M = 9.5$) and nonviolent sexual offenders ($M = 8.0$) scored lowest on the H Scale, whereas patients who had committed “violent property offences” scored highest ($M = 13.8$). There were no differences on the C Scale as a function of index offence.


Summary

This study involved coding the German Version of the HCR-20 on 220 forensic psychiatric patients (209 male). Patients had committed serious offences, been found not criminally responsible, and had been judged to have a high risk for recidivism. Hospitalization is indeterminate; court requires annual progress reports. Mean age of sample was 38.1 (SD = 10.1). Index offences were as follows: homicide (24% of males; 18% of females); assault (21% of males; 27% of females); sexual offences (29% of males; 0% of females); arson (9% of males; 55% of females); property and other offences (15% of males; 0% of females). Diagnostic categories for males were 45% major mental disorder, 35% personality disorder, 20% brain damage, mental retardation or substance abuse disorders. For females, diagnostic categories were 55% major mental disorder, 18% personality disorder, 27% mental retardation. Mean (SD) scores: Total (24.87; 5.90); H (11.97; 3.42); C (5.30; 2.18); R (7.58; 1.86).

Researchers carried out interrater reliability data by having 7 “experienced psychiatrists” rate 50 patients. Cohen’s Kappa for chance-corrected agreement on categorical final risk judgments was .72.

Numerous correlations between H, C, R, PCL:SV, and various inpatient indices of aggression were reported separately for patients with primary diagnoses of major mental disorder versus personality disorder. Correlations between predictors and outcome for patients with major mental disorders were as follows: Minor aggressive acts: threats (HCR-20 Total = .39; H, C, & R = .22, .44, .30; PCL:SV = .30); insults (HCR-20 Total = .30; H, C, & R = .36, .21; PCL:SV = .28). Medium aggressive acts: wilful property damage (HCR-20 Total = .40; H, C, & R = .23, .51, .27; PCL:SV = .24); terror/incitement (HCR-20 Total = .20; H, C, & R = .27, .27; PCL:SV = .21). Major aggressive acts: physical violence toward staff (HCR-20 Total = .23; H, C, & R = .34, .25, .25; PCL:SV = ns); sex offences (HCR-20 Total = .20; H, C, & R = .25; PCL:SV = .21). No measure correlated with fire setting or physical violence toward patients. Correlations for the personality disordered patients were similar for minor aggressive acts, and less consistent for other outcomes.
The researchers concluded that both the HCR-20 and PCL:SV did not predict serious violence consistently. The C-Scale was most consistent for patients with major mental disorders; the PCL:SV for patients with personality disorders alone. Possible reasons include low base rates or small N (neither were reported). The authors claimed that the accuracy of measures for serious violence might have been affected by staff taking measures to prevent violence (hence reducing base rates and likely affecting the behaviour of patients). Staff may have prevented the violence of higher risk patients, hence reducing the correlations between high scores and high incidents of violence.

Four multiple regression analyses were conducted (one for each category of violence as the dependent variable) using the measures’ subscales (i.e., H-scale, C-scale, VRS static, and VRS dynamic) as the predictors. C-scale was the only significant predictor for any institutional incidents and was the most significant predictor for verbal assault. None of the subscales emerged as significant predictors for the outcomes of physical assault and damage to property.

When individual items that comprise the HC composite and VRS scale were considered, those most predictive of inpatient violence were HC composite items that assess a previous diagnosis of mental illness, lack of insight, and active signs of mental illness. Protective factors for institutional violence included VRS items that assess relationship instability, number of young offender convictions, violent lifestyle, and violence throughout the lifespan.

**Project and Scholarly Work**


**Summary**

This study examined retrospectively the predictive ability of the combined HCR-20 H- and C-scales and the Violence Risk Scale 2 (VRS; Wong & Gordon, 2001) within the first six months of admission to a forensic unit. The VRS comprises six static and 20 dynamic factors rated on a 0 (not present/not applicable) to 3 (definitely present/applicable). The measures were completed retrospectively for 44 men using information available at admission. One rater, who was blind to outcome of institutional violence, completed the HC composite and VRS. Another rater, who was blind to risk assessment ratings, rated the incidents of violence. Types of violence coded were physical assault, verbal aggression, and damage to property.

Mean scores on the HC composite were: full scale (19.44, SD = 3.45); H-scale (13.15, SD = 3.25); and C-scale (6.05, SD = 1.98). Total scores on the HC composite and VRS (prorated for omitted items) did not distinguish participants who were aggressive in the institution from those who were not nonaggressive. ROC analyses indicated that the HC and VRS indices, with the exception of the C-scale, tended to not have predictive accuracy for inpatient violence that was greater than chance (the highest value was for the HC composite for physical assaults, AUC = .56, SD = .10). AUC values for the C-scale were larger: any incidents (.72, SD = .08); physical assaults (.60, SD = .11); verbal abuse (.81, SD = .07); and damage to property (.65, SD = .10).

**Henagulph, S., Melvor, R., & Clarke, A. (2012). Risk and recovery group for offenders with mental disorders. Psychiatric Services, 63(1), 94-95.**

**Summary**

The HCR-20 Risk and Recovery Group was developed in the UK as a way to use the HCR-20 to facilitate discussion about risks among mentally disordered offenders. The program promotes transparency about the treating team’s assessment of patient’s risks to others and encourages participants to take responsibility for managing their own risk factors, ultimately reducing their risk. This article provides a short description of this therapeutic program. The authors note that graduates of this program moved to either lower security or were directly discharged to the community.


**Summary**

The current pseudo-prospective study aimed to evaluate risk factors, legal consequences, and recidivism rates for sexual
(i.e., sexual homicide) and nonsexual offending. Psychiatric court reports on 166 men who had committed a sexual homicide between 1945 and 1991 were retrospectively evaluated by three trained forensic psychiatrists and psychologists. The SVR-20, the Static-99, the HCR-20 and the PCL-R were coded based on the information in those reports. PCL-R was used to assess psychopathic syndrome (cutoff score of 20). For statistical group comparisons, cutoff scores of 25 and 20 were chosen for the SVR-20 and HCR-20 respectively. The HCR-20 R subscale and SVR-20 item 19 were not analyzed in this study because they could not be rated with enough confidence for the majority of offenders. Raters were blind to the follow-up data from the federal criminal records. Three types of recidivism were defined: ‘sexual offences’ included rape, sexual assault, sexual child abuse, and sexual homicide; ‘nonsexual violent offences’ were bodily harm, assault, Robbery, kidnapping, nonsexual homicide); and ‘nonviolent offences’ included property offences, possession or trade of illegal drugs, traffic offences, etc.

Interrater reliability was based on 20 reports coded by all three raters. Good IRR was obtained for the PCL-R (ICC = .84 single measure intraclass correlation), the SVR-20 (ICC = .87), that Static-99 (ICC = .84) and the HCR-20 total score (ICC = .77). Follow-up data was available for 139 offenders (83.7% of the original sample). Those without follow-up information had less often committed previous sexual offences before the sexual homicide and less often had high PCL-R and SVR-20 scores. The original sample (N=166) consisted of all Caucasian offenders, 97.6% of whom were German. Twenty-two percent had killed more than one victim, 15.7% committed sexual homicides at two or more distinct occasions and 5.4% were serial killers. The mean age at the time of the first sexual homicide was 26.5 years (SD=8.2; 11.4% were adolescents).

At the time of follow-up based on federal criminal records, 35.5% were still incarcerated in prison or in a forensic psychiatric hospital and 64.7% had been released. The mean time at risk was between 6.4 (any violent offences) to 10.7 years (sexual offences). Of the 90 men released from prison or hospital, 23.1% committed new sexual offences, 18.3% committed new nonsexual violent offences, 35.7% committed any violent offences and 58.4% committed nonviolent offences during the 20 years at risk. The majority of any violent recidivism occurred during the first 5 years after release and sexual recidivism continued over a longer period. While serving their prison sentence for sexual homicide, 10 participants committed new violent offences (5 sexual violence and 5 nonsexual violence). When investigating the influence of different risk factors on the estimated recidivism rates, the analyses were restricted to violent offences. None of the risk assessments or the PCL-R were significantly related to sexual recidivism rates. None of the other risk factors were significant for sexual recidivism. The authors found higher recidivism rates for nonsexual violence in offenders with previous sexual and nonsexual violent offences, in those committing their sexual offences as adolescents and in offenders with higher scores on the PCL-R, HC of the HCR-20, and SVR-20.

Project and Scholarly Works


Summary

This study was a prospective analysis comparing the HCR-20 (Dutch version), BSI and LRA in their ability to assess future risk. The HCR-20 was given before the first supervised leave request and before every extension of unsupervised leave. The BSI was given every half year before treatment evaluation. The LRA was given before extension leave trajectory and advice regarding the extension of the hospital order.

Inter-rater reliability for the HCR-20 (N = 11) was: H-scale (ICC = .92), C-scale (ICC = .91), R-scale (ICC = .95) and Total score (ICC = .98). Inter-rater reliability for the BSI (N = 75) was: Direct aggression (ICC = .84), Obstructionism (ICC = .84) and BSI Risk (ICC = .89). Inter-rater reliability for the LRA (N = 14) was: LRA-SV (ICC = .99) and LRA-DV (ICC = .84).

The distribution of scores for the measures was as follows. With a sample size of 27, the HCR-20 had a mean of 25.2 (SD = 7.54). With a sample size of 62, the BSI had a mean of 4.57 (SD = .37). With a sample size of 16, the LRA-SV had a mean of .6 (SD = 1.72) and the LRA-DV had a mean of −5 (SD = 3.34).

Correlations between the HCR-20 and the BSI were conducted. The HCR-20 total score was correlated with BSI-Obstructionism (r = -.38; p < .1), BSI-Direct aggression (r = -.37; p < .1) and with BSI-Risk (r = -.43; p < .05). The HCR-20 H-scale was correlated with BSI-Direct aggression (r = -.35; p < .1) and with BSI-Risk (r = -.36; p < .1). The HCR-20 C-scale was correlated with BSI-Obstructionism (r = -.51; p < .01), BSI-Direct aggression (r = -.46; p < .05) and with BSI-Risk (r = -.56; p < .01). The HCR-20 R-scale was not correlated with any BSI score. The HCR-20 total score was correlated with LRA-SV scale (r = .77; p < .01) and with the LRA-DV scale (r = .5; p <
The HCR-20 H-scale was correlated with LRA-SV scale ($r = .73; p < .01$) but not with the LRA-DV scale. The HCR-20 C-scale score was correlated with LRA-SV scale ($r = .7; p < .01$) and with the LRA-DV scale ($r = .52; p < .1$). The HCR-20 R-scale was correlated with LRA-SV scale ($r = .58; p < .05$) but not with the LRA-DV scale.


**Summary**

The Leave Risk Assessment (LRA) is an actuarial risk assessment tool composed of both historical and treatment-related subscales, developed to assess the risk of serious reoffending by forensic psychiatric patients. In this retrospective study the psychometric properties of the LCA were examined. As part of this examination, the association between LRA and the HCR-20 was also investigated. The sample used in this study was drawn from the same population on which the LRA was developed. This population included 78 TBS (i.e., Dutch forensic) patients who committed a violent or felony offense during authorized leave between 1997 and 2003. Additionally, 117 non-offenders were randomly selected from TBS patients who were on authorized leave and terminated TBS between 1998 and 2003.

The LRA and HCR-20 were coded using information from the TBS patient files at the Ministry of Justice. Raters were blind to which group (reoffender vs. non-reoffender) patients were in. Agreement between raters was calculated on the basis of 20 cases. The inter-rater reliability of the LRA-HIS was excellent (ICC = .93) and good for the LRA-TRIS (ICC = .62) and the LRA total score (ICC = .72). The inter-rater reliability was also good for the HCR-20 total score (ICC = .74) and the SPJ summary risk rating of high, moderate, or low risk (ICC = .61). Serious criminal offenses were operationalized as offenses with a maximum prison sentence of at least four years and were of a violent/and or sexual nature. Of the 78 patients who committed an offense during leave, 63.3% committed a serious offense, and 36.7% committed a less serious offense.

Concurrent validity analyses indicate that the LRA-HIS was moderately correlated with the historical scale of the HCR-20 ($r_s = .46, p < .001$) and had smaller correlations with the clinical and risk management scales of the HCR-20 ($r_s = .20$ for each, $p < .01$). The LRA-TRIS was strongly associated with the C and R scales, and the HCR-20 total score ($r_s = .58, r_s = .56$ and $r_s = .60$, respectively, all $p < .001$), but less so with the Historical items of the HCR-20 ($r_s = .27, p < .001$). Finally, the LRA total score correlated strongly with the HCR-20 total score ($r_s = .57, p < .001$).

AUC values for LRA and HCR-20 total and subscale scores were reported for serious and general offending. The LRA-HIS was not significantly different from the historical scale score of the HCR-20 in assessing reoffending during leave. In contrast, the LRA-TRIS assessed risk of serious and general reoffending during leave significantly better than the clinical and risk management scales of the HCR-20 ($p < .05$ and $p < .01$, respectively). Moreover, the LRA total score had a significantly higher predictive validity compared to the HCR-20 total score ($p < .001$) and HCR-20 summary risk judgment ($p < .001$). With respect to serious offending, AUCs were .76, .75, and .84, for LRA historical, treatment related and total scores, respectively, and .69, .68, .66, .62 and .69 for HCR-20 total, historical, clinical, risk management scales, and final risk judgments, respectively. With respect to general offending, AUC values were 0.77, 0.73 and 0.83 LRA historical, treatment related and total scores, respectively, and 0.70, 0.70, 0.64, 0.62 and 0.065 for total, H, C, R and final risk judgment, respectively.

Multivariate logistic regression analyses were conducted with the HCR-20 total score, the HCR-20 summary risk judgment and the LRA total score using the forward stepwise method. The LRA total score produced a significant fit for serious reoffending, $\chi^2(1, N = 195) = 57.41, p < .001$, and general reoffending, $\chi^2(1, N = 195) = 66.53, p < .001$. In both analyses, the HCR-20 total score and the final risk judgment did not produce a significant improvement to the model after the LRA was entered. Subsequently, the authors tested the incremental validity of the Clinical and Risk management scales of the HCR-20 on the LRA-TRIS, using the same procedure. The LRA-TRIS produced a significant fit ($\chi^2(1, N = 195) = 38.27, p < .001$) and $\chi^2(1, N = 195) = 33.96, p < .001$ for serious and general reoffending, respectively, the clinical ($p = .81, p = .99$, respectively) and risk management scales ($p = .59, p = .88$, respectively) of the HCR-20 were not entered into the model because they did not contribute significantly to the improvement of the model.

The authors concluded that the LRA can have significant contribution in the decision-making process regarding authorized leave. Although the authors concluded that the LRA had incremental predictive value over the HCR-20, they noted that because the HCR-20 was used as an actuarial instrument (by summing scores) rather than an SPJ tool and that the psychopathy item was not included results may underestimate the true utility of the tool.

**Project and Scholarly Works**

Summary

Many risk assessment instruments are available for assessing an individual’s risk of violence; however, few validated tools are available to assess risk of suicide. The present study was the first to investigate the reliability and concurrent validity of the Suicide Risk Assessment and Management Manual (S-RAMM). Participants were inpatients at the only forensic psychiatric unit in Ireland. The authors report acceptable levels of internal consistency and interrater reliability for the S-RAMM, based on coding 25 cases.

The HCR-20 was then used to establish concurrent validity of the S-RAMM based on independent rating made on 81 patients. HCR-20 total scores and S-RAMM total scores were related \( r = .48 \). The H scale did not correlate with the background items on the S-RAMM \( r = .11, p = .34 \). The C and R scales correlated well with the corresponding Current and Future subscales on the S-RAMM \( r = .50 \) and \( r = .44 \), for the C and R scales, respectively. The dynamic totals also correlated well \( r = .62 \).

Next, it was investigated whether the S-RAMM and HCR-20 stratified across the levels of security within the hospital. The S-RAMM total score \( F = 11.3, p < .001 \), background subscale \( F = 4.3, p = .001 \), current subscale \( F = 7.6, p < .001 \), and future subscale \( F = 10.1, p < .001 \) all differed significantly across the six levels of security. A similar pattern was also seen for the HCR-20 total score and subscale scores, as they stratified across the security levels.

The authors discussed the implications of the results and the need for future research regarding the use of structured assessment tool for risk of suicide.


Summary

In a second part to the S-RAMM validation study reported above, the authors report the results of the prospective portion of the study. Participants included the 81 inpatients that were report above. Participants were followed for a mean number of 184.8 days, and incidents of self harm, suicide and violence were recorded.

The S-RAMM was found to be an excellent predictor of self harm \( AUC = .902 \), suicidal ideation \( AUC = .875 \), and violence \( AUCs = .744 - .790 \). The HCR-20 was also predictive of violence as defined by the HCR-20 \( AUCs = .766, .775, \) and \( .796 \), for the H scale, combined C and R scales, and total score, respectively. The HCR-20 was also predictive of violence as defined more broadly to include aggression against objects \( AUCs = .738, .728, \) and \( .760 \), for the H scale, combined C and R scales, and total score, respectively. In addition, the HCR-20 was found to be predictive of both self harm incidents \( AUCs = .776, .910, \) and \( .887 \), for the H scale, combined C and R scales, and total score, respectively and suicidal ideation \( AUCs = .553, .754, \) and \( .705 \), for the H scale, combined C and R scales, and total score, respectively). The authors also report predictive validity analyses for the GAF and PANSS.

HCR-20 Review and Annotated Bibliography


Summary

Current consensus in England supports the use of Structured Professional Judgment (SPJ) instruments; however, the extent to which these instruments are actually used in practice is unknown. This study attempted to determine which tools are used in medium secure forensic psychiatric units and to measure the perceived value of the tools.

A total of 47 medium secure units for adults were sent a questionnaire asking for details of the unit and specifically which violence risk instruments were used at the facility with a list of named instruments provided. Survey respondents were asked to indicate whether they used each of the instruments “frequently,” “occasionally,” or “never” and were subsequently asked to rate the utility of the instruments they endorsed using on a 5 point likert scale. Of the 47 units sent the questionnaire, responses were received from 29 (a 62% response rate).

The HCR-20 and PCL-R were the most widely used instruments, and they were often used in conjunction. The START was rarely used but received the highest ratings in terms of its utility. With regards to risk for sexual violence, the RM2000, Static-99 and SORAG were reported to be used the most frequently.

Qualitative interviews were conducted with several sites subsequent to the questionnaire. Of those who reported...
using the HCR-20, common reasons for its use included it being “accessible to all disciplines; it provided comprehensive information about violence risk; it helped with risk management; it was tailored to the individual because it included specific risk scenarios; its dynamic content allowed monitoring of change; and it was widely understood by other clinicians.”

**Project and Scholarly Works**


**Summary**

The Dangerous and Severe Personality Disorder (DSPD) initiative in England and Wales provides specialized care to high risk personality disordered individuals in prison and secure psychiatric facilities. In this dissertation, the predictive utility of several risk assessment tools were prospectively examined for institutional infractions among personality disordered DSPD patients in a high security forensic psychiatric setting. There were two studies: in the first study, the predictive utility of the HCR-20, Violence Risk Scale (VRS), Static-99, and Risk Matrix 2000 Scale were examined. In the second study, the predictive utility of the PCL-R and International Personality Disorder Examination were examined.

The sample used in this study was comprised of 44 males who were an average of 34.41 years of age (SD = 8.47). The most common diagnoses in the sample were alcohol and substance-related disorders (57% and 52%, respectively), post-traumatic stress disorder (36%), and major depressive disorder (24%). Using the IPDE, the most common personality disorder (PD) diagnoses were for Antisocial PD (73%), Borderline PD (41%), and Narcissistic PD (16%). Using the recommended UK cut-off score of 28 on the PCL-R, 49% of the sample met criteria for psychopathy. All patients had been charged with or convicted of a criminal offence, and 89% had committed one or more violent offences. Incidents of aggression, including damage to property, verbal aggression, and interpersonal physical aggression directed to staff or other patients were identified over a 12 month period or the date of transfer of the patient to another unit. Of the sample, 38% of patients damaged property on at least one occasion and 38.6 % engaged in physical aggression toward another person on at least one occasion. Due to the fact that almost all of the patients were verbally aggressive during the follow-up period only the accuracy of the assessment measures in predicting incidents involving damage to property and interpersonal physical aggression were examined.

Correlations were reported for each of the component and total scores of the HCR-20, VRS and PCL-R. HCR-20 SPJ ratings were significantly correlated with the VRS-D ($r = 0.71$, $p < .001$), PCL-R Factor 1 ($r = 0.48$, $p < .05$) and PCL-R Facet 2 scores ($r = 0.57$, $p < .01$). HCR-20 total scores were significantly correlated with VRS total ($r = 0.50$, $p < .01$) and VRS-D ($r = 0.59$, $p < .001$) scores, as well as PCL-R Total ($r = 0.54$, $p < .001$), PCL-R Factor 2 ($r = 0.32$, $p < .05$), Facet 1 ($r = 0.36$, $p < .01$) and Facet 2 scores ($r = 0.56$, $p < .001$).

AUC values for the full follow-up period and the initial 12 months were reported for each measure for incidents involving damage to property and incidents involving interpersonal physical aggression. Among the risk assessment tools, the HCR-20 Total, Clinical, and Risk Management scales as well as the SPJ showed at least moderate levels of predictive accuracy for both outcomes for at least one of the two follow-up periods. For both damage to property and interpersonal physical aggression outcomes, the HCR-20 SFRJ was found to have a moderate to large level of accuracy, with AUCs between .73 to .80 (all $ps < .05$). For damage to property in the initial 12-month follow-up period the effect size was in the medium to large range although the AUC failed to reach statistical significance, likely due to low power. Using the full follow-up, the HCR-20 Total score also showed a moderate level of accuracy for both outcomes (AUCs = 0.68 and 0.70, for interpersonal aggression and aggression towards property respectively, both $ps < .05$). The Clinical and Risk Management scales similarly predicted both outcomes (although the pattern of significant AUCs according to outcome and follow-up period varied between the two and only the Risk Management scale predicted interpersonal physical aggression for both follow-up periods) with effect sizes generally falling in the medium to large range for all these components. The HCR-20 Historical scale failed to predict the outcomes for either follow-up period with small effect sizes (AUCs between .48 and .58). In contrast to the HCR-20, performances for the VRS (Total score and Static and Dynamic Factor scales), the RM 2000 and the Static-99 were all less consistent and generally weaker. Total scores for these tools did predict damage to property in the first year following admission but no total or scale scores significantly predicted interpersonal physical aggression. The author noted that effect sizes varied and those few in the medium range do suggest that additional research with larger samples is required before conclusions are drawn about the predictive utility of these tools within the DSPD service. It should be underscored that with an N of 44, power would have been inadequate to detect some moderate effect sizes.
Summary

Many risk assessment instruments are used in assessing offenders entering the Dangerous and Severe Personality Disorder (DSPD) service in England and Wales. As a result, the present study investigated the utility of various violence risk assessment instruments on this specialized population. This unit provides focused care for high risk individuals with personality disorders. It is routine practice for initial assessments of individuals entering this unit to include structured assessments using the HCR-20, the Violence Risk Scale (VRS), the RM 2000, the Static-99, and the PCL-R. The ability of these instruments to predict institutional infractions among this population has not previously been established.

The present study focused on the first 51 consecutive admissions to the unit, yet full data was only available on 44 participants. All the participants were male and mostly Caucasian with a mean age of 34.41 years (SD = 8.47). The most common Axis I disorders included substance related disorders, posttraumatic stress disorder, and major depressive disorder. Antisocial personality disorder was the most common personality disorder. The participants were followed for an average of 570.3 days following their assessments. Incidents of damage to property and interpersonal physical aggression were recorded.

Many of the risk assessment measures correlated with one another. The HCR-20 final risk judgments were related to the total scores ($r = .69$), H scale ($r = .44$), and R scale ($r = .70$), but not the C scale ($r = .29$, ns), although this appears to be attributable to low power. The total score correlated with all of the scale scores ($r = .73$, .71, and .72, for the H, C, and R scales, respectively). The H scale correlated with the C scale ($r = .36$), but not the R scale ($r = .28$, ns; although note power problems). Finally, the C and R scales were correlated ($r = .35$). The HCR-20 final risk judgments also correlated with the VRS D items ($r = .71$) and Factor 1 of the PCL:R ($r = .48$). The HCR-20 total scores also correlated with the VRS total scores ($r = .50$) and D items ($r = .59$), as well as the PCL:R total scores ($r = .54$), Factor 1 ($r = .50$), and Factor 2 ($r = .32$). The H scale items were also correlated with the VRS total scores ($r = .51$) and D items ($r = .58$), as well as the PCL:R total scores ($r = .62$), Factor 1 ($r = .34$), and Factor 2 ($r = .52$). The C scale was correlated with the VRS total scores ($r = .40$) and D items ($r = .39$). The R scale was correlated with the VRS D items ($r = .45$) and Factor 1 of the PCL:R ($r = .35$). Several other correlations among these instruments were also reported.

With respect to predicting damage to property, several instruments were found to be predictive including the HCR-20 final risk judgments (AUC = .73), the HCR-20 total scores (AUC = .70), the R scale (AUC = .77), the VRS total scores (AUC = .72), the VRS S scale (AUC = .67, and the RM 2000V (AUC = .74). Point biserial correlations and correlations between the frequency of this outcome and assessments scores confirmed these findings. The HCR-20 final risk judgments were also the only feature to predict repetitive physical aggression (AUC = .74). The final risk judgments and the R scale were also correlated with the number of days to committing physical aggression ($r = -.48$ and -.33, respectively). The results were discussed in terms of the utility of these instruments for the DSPD population.

Summary

The purpose of the present study was to gather reliability data on assessments of patients who were considered potentially suitable for the specialist dangerous and severe personality disorder (DSPD) programme in the United Kingdom. A random sample of patients ($N = 32$) at Rampton high security hospital who had been assessed by clinical staff (usually psychologists) at the hospital between April 1997 and November 2002 were selected. Two participants already had left the hospital, which yielded a final sample size of 30 men.

The data compared for reliability analyses were scores on the HCR-20, PCL-R (Total, Factor 1, and Factor 2), and each DSM personality dimension on the International Personality Disorder Examination (IPDE). With respect to the original ratings, it was unusual for a single assessor to complete the entire assessment. With respect to the research
ratings, a single psychologist completed all ratings between February 2002 and April 2002. Another researcher who was blind to the research interview data obtained details of the original assessments from records.

Of the 30 participants, 11 refused to participate in an interview and one other was not contacted for other reasons. Participants who were interviewed were similar to participants who were not interviewed in terms of mean duration of admission (120 months vs. 177 months for interviewees vs. refusers, respectively) and mean age (38 years vs. 45 years for interviewees vs. refusers, respectively). Of the interviewed men, 12 had IPDE ratings on both occasions, 15 had PCL-R and HCR-20 ratings, and one did not have any previous assessment identified. There was substantial variation in time intervals between the original and research assessments: HCR-20 (Md = 15 months, range = 1-25 months); PCL-R (Md = 11 months, range = 6-60 months); IPDE (Md = 9 months, range = 1-18 months).

Intra-class correlation coefficients (ICC1) for all indices examined ranged from 0.38 to 0.73. For the HCR-20 total score, ICC1 was 0.57, p < .01. Values for the PCL-R indices were as follows: Total (ICC1 = 0.59, p < .01); Factor 1 (ICC1 = 0.49, p < .05); Factor 2 (ICC1 = 0.44, p < .05). All values for the HCR-20 and PLC-R indices were based on 15 participants and would be considered to represent fair levels of clinical significance according to criteria outlined by Cicchetti and Sparrow (1981).

Rater bias was assessed separately from intra-class correlation coefficients so that systematic differences in scoring between raters could be identified independently of agreement. Significant rater bias was observed for the HCR-20 total score and the IPDE antisocial and avoidant scales, but not for any PCL-R index. The significant F-ratios for rater bias were as follows: HCR-20 (F = 13.1, p < 0.01); IPDE antisocial (F = 9.0, p < 0.05); IPDE avoidant (F = 5.7, p < 0.05). Higher scores were recorded at the second assessment for the HCR-20 (26.8 vs. 22.9), for the IPDE antisocial scale (19.0 vs. 13.9), and for the IPDE avoidant scale (3.6 vs. 1.7), which makes it unlikely that improvement in clinical state would explain the rater bias.

The authors concluded that their results supported the need for better training in the use of standardised instruments.

### Summary

The purpose of the present study was to determine if measures of anger, impulsivity and mental health symptoms would improve the predictive validity of the HCR-20 or VRAG for institutional aggression. The authors also sought to determine subtypes of patients who are violent in institutions. This study was a prospective study conducted at a long-term psychiatric hospital in California were approximately 80% of patients are under a forensic commitment. The sample comprised 154 patients post-trial between July 2002 and September 2005. Of those, 108 completed the required assessments.

The overall sample was mostly male (84%) and Caucasian (72%). The modal commitment offences were assault and/or battery (39%) followed by murder/manslaughter (24%). Most participants were committed under the NGRI statute and were diagnosed with schizophrenia (53%), or schizoaffective disorder (19%), with the remainder diagnosed with mood disorders, substance use disorders or other disorders. The average age of participants was 45.6 years with an average length of stay in the hospital of 5.9 years. All violence risk assessments were coded by trained doctoral level psychologists. Inter-rater reliability ranged from an average intraclass correlation of .86 for the HCR-20 R subscale to .97 for the total PCL-R score. Routine recalibrations were performed to prevent rater drift. The outcome was coded from Special Incident Reports (SIRs) which were completed for incidents of physical aggression (against either patient or staff), verbal aggression (against either patient or staff), self-injurious behaviour, property damage, unauthorized absences, fire-setting, and other categories related to staff behaviour.

The average length of follow-up for the sample was 2.48 years (SD = .88, .97– 4.01 years). The average rates of physically aggressive acts per year were .11 (SD = .34) for staff-directed aggression, .16 (SD = .40) for patient-directed aggression, and .28 (SD = .64) for both categories combined. SIRS involving verbal aggression were very rare so the rest of the analyses focused on physical aggression. The percentage of patients with one or more aggressive incidents was 16% (staff-directed), 22% (patient-directed), and 28% (combined). The means of the risk assessments were as follows: HCR-20 Total M = 23.76, SD = 6.22; VRAG M = 5.36, SD = 9.89; PCL-R M = 16.18, SD = 7.90.

In terms of aggressive incidents in total (staff and patient-directed), the HCR-20 was the only significant predictor (AUC = .65) and the R subscale was the strongest predictor (AUC = .66). For staff-directed aggression, the relationships

### Project and Scholarly Work

were weaker but the HCR-20 total score was still a significant predictor (AUC = .65) as well as Factor 2 of the PCL-R (AUC = .66) and the VRAG with and without the PCL-R item (AUC = .65, .67). For patient-directed aggressive acts the C and R subscale (AUCs = .66, .70) evidenced significant moderate relationships as well as Facet 1 of the PCL-R (AUC = .65). Although the HCR-20 had the strongest relationship to the outcomes, the H subscale was unrelated to any type of aggression. Next, the authors used logistic regression to partial out the unique variance for each predictor and then ROC analyses were conducted again. Plus the PCL-R was removed from the VRAG and HCR-20. When the PCL-R scores were controlled for, the VRAG continued to be a modest predictor of staff-directed aggression and the HCR-20 was also still significant for total aggressive incidents and patient-directed aggression but not the staff-directed aggression. This same pattern emerged for the R-subscale. Given that the other predictors (BPRS, Novaco, and Impulsivity) were no longer significant once the R subscale was controlled for in analyses, the authors concluded the extra measures would not add substantially to the prediction of institutional aggression. Using cluster analyses, the authors identified 4 clusters which differed in terms of the means on the risk assessments as well as types of institutional aggression.

Project and Scholarly Work


Summary

Research examining the predictive validity of risk assessment instruments for predicting inpatient violence in secure forensic psychiatric units was claimed by the authors to be limited. As such, the present study sought to examine whether various instruments were predictive of different types of violence in an inpatient setting. A total of 238 participants housed in a long-term forensic psychiatric hospital were involved in this study. The participants were mostly male (86%) and Caucasian (63%) with an average age of 46.6 years. Many of the participants had diagnoses of schizophrenia (45%) or schizoaffective disorder (27%). The mean length of stay prior to the study period was 4.5 years. The mean follow up length following baseline assessment was 2.52 years (SD = 1.55 years). Numerous instruments were used: the HCR-20, PCL-R, BPRS, Barratt Impulsivity Scale (BIS), and the Novaco Anger Scale and Provocation Index (NAS-PI). A total of 25% of the participants engaged in some type of violent behaviour during the follow up period. Of these aggressive acts, 58% were classified as impulsive, 14% as predatory, and 15% as psychotic.

Several group differences between aggressive and non-aggressive individuals were found. Non-aggressive individuals had lower HCR-20 total scores, C scale scores and R scale scores, as well as lower scores on the BPRS (total score and subscales).

With regard to impulsive physical aggression, the HCR-20 total score, C scale, and R scale were predictive (AUCs = .67, .69, and .66, respectively). Other instruments that were predictive of this type of aggression included the PCL-R (AUC = .64) and the BPRS total score (AUC = .70).

With regard to predatory physical aggression, the HCR-20 total score, C scale and R scale were predictive once again (AUCs = .68, .68, and .69, respectively). Other instruments that were predictive for this type of aggression included the PCL-R (AUC = .66) and the NAS total score (AUC = .67).

With respect to psychotic physical aggression, the C scale was predictive (AUC = .66), as was the BPRS total score (AUC = .77) and all of the BPRS subscale scores.

The authors also report the predictive ability of the same instruments across the same three types of physical aggression for only the first 12 months of the follow up period. The same overall patterns are evident for the short term prediction of these outcomes. The main difference is that the Novaco subscales are much more predictive of impulsive and predatory aggression in the short term.

The results are discussed in context of predicting and managing inpatient aggression on secure forensic units.

Project and Scholarly Work


Summary

The authors examined retrospectively the predictive validity of the Historical and Clinical scales for inpatient violence. Participants were 21 women and 74 men who had been admitted sequentially to a medium secure unit and residing for at least four months. Participants’ mean age was 35 (range: 18-62).
Data were coded for the Historical and Clinical scales by two raters using information that would have been available in the first two weeks after admission. The authors cited insufficient variance and poor interrater reliability as reasons for not completing the Risk Management scale. The Historical scale was completed on the basis of medical reports available at admission. Items on the Clinical scale were rated from nursing observation notes regarding the behaviour and clinical state of the individual during the first two weeks post-admission.

Inpatient violence was defined as acts of physical aggression towards a person or property. Violent episodes were coded from a database of critical incidents recorded by nursing staff. Length of follow-up was not reported.

Mean scores on the Historical and Clinical scales were 9.0 \((SD = 3.0)\) and 5.5 \((SD = 2.4)\), respectively. Correlations and AUC values for violence for the combined Historical + Clinical (HC), Historical (H), and Clinical (C) scales, respectively, were: HC \((r = .49, p < .01; \text{AUC} = .65, p = .03)\); H \((r = .14, p = .06 \text{AUC} = .55, p = .50)\); C \((r = .40, p < .01; \text{AUC} = .68, p = .01)\). In terms of the predictive power of the individual items, none of the H scale items except H10 Prior Supervision Failure were statistically significant. All C scales items correlated significantly with violence, with the largest Rho observed for C4 Impulsivity \((.55, p < .01)\) and the smallest Rho observed for C5 Unresponsive to Treatment \((.18, p = .04)\). AUC values for the C items were: C1 Lack of Insight \((.55, p = .50)\); C2 Negative Attitudes \((.66, p = .02)\); C3 Active Symptoms of Mental Illness \((.60, p = .15)\); C4 Impulsivity \((.77, p = .01)\); C5 Unresponsive to Treatment \((.54, p = .61)\).

The authors divided the sample in four groups: those not violent, those violent fewer than five times during their admission, those violent between five and ten times, and those violent more than ten times. Predictive power of the HC, H, and C scales increased as a function of frequency of violence observed.

**Background:** A substantial number of forensic psychiatric patients also show aggressive behaviour while being admitted to a psychiatric hospital. Risk assessment can therefore be of importance not only for estimating the risk of recidivism after treatment, but can also be used to protect the hospital staff. **Aim:** To find out to what extent scores on the Historical Clinical Risk Management-20 (HCR-20) can predict aggressive behaviour during inpatient treatment in a forensic psychiatric department. **Method:** In total, 102 patients were included in our analysis. Of these, 43 patients had caused 174 aggressive incidents between January 2005 and August 2008. The incidents were recorded by staff members who used the Staff Observation Aggression Scale-Revised (SOAS-R). **Results:** On average, the group of patients involved in one or more aggressive incidents were found to have higher HCR-scores than patients who were not involved in aggressive incidents. The area under the curve (AUC)-value of the HCR-20 total score was 0.70 as far as the prediction of aggression was concerned. Logistic regression analysis suggested that particularly the C-subscale items assessing impulsivity (item C4) and the patient’s response to treatment (item C5) may be able to predict aggressive behaviour fairly accurately in a particular department. **Conclusion:** The HCR-20 can predict to a certain extent which patients will engage in violent behaviour while receiving treatment in a forensic psychiatric department. These results correspond to those of similar earlier investigations which showed that the HCR-20 could predict that patients would engage in further violent or criminal behaviour after being discharged from hospital.

**Project and Scholarly Work**


**Summary**

The author assessed whether Theory of Mind (ToM) deficits among 30 male schizophrenics in high security psychiatric care were related to HCR-20 scores and assessments of clinical outcome three years after the initial ToM assessments. Most patients had prior criminal histories for violent offences including rape, indecent assault, manslaughter, assault with bodily harm, grave bodily harm, and arson. The majority of patients also had histories of substance misuse. All non-social cognitive and ToM measures were assessed as part of a routine neurological assessment at admissions. The clinical outcome measures,
including the HCR-20 were completed three years later by trained clinicians.

The mean scores for the HCR-20 are as follows: H subscale $M = 13.5$, $SD = 2.5$, 9–18; C subscale $M = 4.9$, $SD = 2.5$, 0–10; R subscale $M = 6.1$, $SD = 2.3$, 1–10; HCR-20 total $M = 24.5$, $SD = 5.5$, 13–37. The second-order Modified Advanced Theory of Mind Test (MAT) was significantly correlated with HCR-20 R subscale ($r = .42$). The Revised Eye Test (RET) was significantly correlated with HCR H subscale ($r = .46$), the R subscale ($r = .48$) and the HCR-20 Total ($r = .49$). When controlling for the WAIS only the relationship with the R subscale remained. The WAIS F IQ was significantly correlated with the H subscale ($r = .37$) and the R subscale ($r = .42$). The results suggested that many dimensions of neuropsychological function are related to risk for violence.

Project and Scholarly Work


Summary

The authors argued that while many studies demonstrate good predictive utility of the HCR-20 for violence, less attention has been paid to its clinical applicability. The present study was a true prospective study into the utilization of the HCR-20 as part of a clinical routine at a forensic psychiatric unit in Denmark. The final study sample consisted of 81 male patients discharged from the unit between 2006 and 2007 to other psychiatric hospital wards (38%), prison wards (7%) or the community (48%). All data were collected and scored by treating clinicians as part of the clinical routine of the unit. Risk of future violence was assessed by the HCR-20 and psychopathy was scored using the PCL: SV. The outcome variables of interest to the study were inpatient aggression and post-release recidivism. Aggression during hospitalization was recorded on the revised version of the Staff Observation Aggression Scale (SOAS-R), which was also implemented as part of the routine at the unit. Aggression included any verbal, non-verbal or physical behavior that was threatening, or physical behavior that actually did cause harm. New convictions post-discharge were collected from date of discharge until November 2008 using information extracted from the Danish National Crime Registry. The mean follow-up time was 21 months ($SD = 6.28$). Recidivisms were categorized into the following categories: any violent crime, any non-violent crime or any crime (an omnibus category that included all crime).

The mean age of the patients was 35.7 years ($SD = 10.49$). Seventy-nine percent were diagnosed with schizophrenia, 3% had affective disorders, 9% had personality disorders and 6% had other diagnoses. Of the sample, 68% had a substance use disorder. Mean admission time for the index hospitalization was 432 days. Approximately 37% of the patients had one or more aggressive incidents during the index hospitalization. A total of 43% of the sample was reconvicted during the follow-up and 23% were reconvicted for a violent crime. There were no significant differences in recidivism rates between patients discharged to another institution or to the community.

Mean scores on the HCR-20 were $13.73$ ($SD = 3.48$), $5.42$ ($SD = 2.48$), $5.88$ ($SD = 2.18$) and $25.05$ ($SD = 6.86$) for the historical, clinical, risk management and total scales. Patients with aggressive inpatient behavior had significantly higher scores on the HCR-20 total scores (mean $28.3$ vs. $23.8$), H scale (mean 15 vs. 13.1) C scale (mean 6.6 vs. 5.1), R scale (mean 6.7 vs. 5.6) and structured final risk judgment (mean 2.1 vs. 1.7). In regards to violent reconvictions, the predictive ability of the HCR-20 total score was in the moderate range with an AUC value of .66. The AUC values were .68, .62 and .58 for the historical, clinical and risk management scales.

The structured final risk judgment showed poor predictive accuracy (AUC = .56). Only the HCR-20 total score and historical scale were statistically significant. With regards to inpatient aggression, the predictive ability of the HCR-20 was in the moderate range for the structured final risk judgment (AUC = .64). The AUC values were 0.70, 0.68, 0.66 and 0.66 for the total, historical, clinical, and risk management scales of the HCR-20. All scales, except the structured final risk judgment, were statistically significant in predicting future inpatient aggression.

While the predictive validity of the HCR-20 was lower compared with previous findings, the authors note that this was likely the result of the HCR-20 being used to guide risk management strategies which reduced subsequent violent behavior. Further the authors’ note that because recidivism was recorded as a new criminal conviction only underreporting may have occurred. Similarly, they note that aggression during hospitalization may have been underreported as only the SOAS-R was used and not collateral file information. These limitations may have also affected the predictive accuracy of the measure.

See Also

Project and Scholarly Work


Summary

This study investigated the interrater reliability of the HCR-20 among 21 practicing clinicians (12 males and 9 females) in a large forensic psychiatric program in Eastern Canada. The aim of this study was to determine whether acceptable levels of reliability were obtained when the HCR-20 was scored by practicing clinicians with varying levels of training, experience, and familiarity with the SPJ model. Each rater independently scored the HCR-20 for three cases designed to vary in overall risk level and completed a questionnaire following the ratings. The questionnaire asked about the rater’s gender, education, prior training in the HCR-20 and other forensic assessment instruments, and years of experience in forensic mental health.

For each case, HCR-20 total and subscale means were provided as a function of rater gender (male, female), profession (psychiatry, psychology), expertise (prior HCR-20 training, no prior training), and experience (more than 10 years, less than 10 years). Overall, raters with varying professional background, training, and experience generated comparable total and subscale scores across the three cases. For one case, participants with prior HCR-20 training had significantly higher ratings on HCR-20 total score than participants with no prior HCR-20 training. For another case, participants with more than 10 years of experience had significantly higher ratings on HCR-20 total and H scores than participants with less than 10 years of training. No other significant differences in ratings were observed.

The authors provided ICC values separately for each rater characteristic. ICC values ranged between .89 and .94 for the H scale, .77 and .95 for the C scale, .58 and .75 for the R scale, and .91 and .94 for the total scale. For the sample as a whole, ICC values were .92, .86, .65 and .92, for the H, C, R and Total scales, respectively. For the summary risk ratings, half of all kappa values showed substantial to excellent agreement, while a further 20% fell into the moderate range. Case one was rated moderate risk by 83% of raters, case two was rated high risk by 67% of raters, and case three was rated low risk by 96% of raters. The authors concluded that raters from different disciplines and with varying levels of training and experience can achieve good to excellent levels of reliability when scoring the HCR-20.

See Also


Project and Scholarly Works


Summary

This was a retrospective validation study in the Netherlands of the Dutch version of the HCR-20 (Philipse, de Ruiter, Hildebrandle & Bauman, 2000). This research study used subset of 69 patients from three hospitals from a larger prospective study on assessing risk for re-offending. The research was conducted without using Item 7 (Psychopathy) from the H scale. The sample consisted of 64 males and 5 females. The types of offenses were categorized as violent, sex, and arson. Patients had left the hospital between 1/1/96 and 12/31/98. Re-offending data was collected on 1/22/02 with an average of 4 years and 4 months of time for patients to have been outside of the hospital. 21 (30%) had been found to have had renewed contact with the law.

Inter-rater reliability for the HCR-20 (Dutch version) was:

- ICC HCR-20 Total = .90, ICC H-scale = .79, ICC C-scale = .76, ICC R-scale = .67. The total and R-scale scores were significantly lower for patients discharged from the hospital according to hospital advice. R-scores were predictive of type of discharge (AUC = .67). HCR-20 (Dutch version) was most effective for non-sexual offenders. Deleting females did not alter the findings. The postdictive validity AUCs for committing a violent act (when excluding sex offenders) were: HCR-20 total score = .67, H-scale = .72, C-scale = .60, R-scale = .58. The postdictive validity AUC for clinical judgment was .64, as was the number of previous convictions. Reducing the HCR-20 into smaller units increased the postdictive AUC values. Using only the H2,
H4, H5, H10, C3 and C4 items achieved an AUC of .82. Using only the H2, H5, H10 and the C4 items achieved an AUC of .90.

In terms of decision making, the 4-item version of the HCR-20 with a cut-off of 50% identified all offenders with 2.2 false positives per true positives. The 4-item version of the HCR-20 with a cut-off of 80% identified 5 of 8 offenders with .6 false positives per true positives. Implications for the clinical assessment of risk of re-offending and the best composition of the HCR-20 items are discussed.


Summary

Risk assessment in Dutch forensic psychiatry (tbs) is still dominated by an unstructured clinical approach. Researchers have argued in favour of a standardised approach because international research reports limited predictive validity of clinical approaches. The Dutch version of the clinical-actuarial debate is briefly summarised in this article. A study is presented that evaluates the validity of an international risk assessment tool, the HCR-20 in tbs. This shows that using the HCR-20 may improve risk assessment under certain conditions, although unstructured clinical judgement performs quite well too. Also, it is shown that clinically adjusted HCR-scores are slightly better than actuarial scores. However, in the final analysis historical predictors outperform all other measures. It is concluded that the HCR-20 may constitute a meaningful addition to Dutch risk assessment practice, though it is imperative that all persons dealing with this and similar instruments have a clear view of their limitations.

Combined seclusion data across first and second years of hospitalization showed that for almost half of the sample (42%) seclusion was necessary. Similarly in years one and two, the greatest proportion of patients were secluded either once (13%) or twice (11%). The proportion of patients that required seclusion on three or more occasions revealed a declining linear trend ranging from 5% to 1%. Neither gender nor age correlated with any seclusion variables, with the exception of age having a negative but mild association with average seclusion duration (r = -0.29, p < .05). Mean scores in the sample were: PCL-R Total M = 12.50 (SD = 6.81); Factor 1 M = 3.49 (SD = 2.93); Factor 2 M = 7.92 (SD = 3.82); VRAG Total M = -1.04 (SD = 10.69), VRAG Bin M = 4.47 (SD = 1.56); LSI-R M = 24.06 (SD = 7.40); HCR-20 Total M = 24.44 (SD = 7.17), H M = 11.90 (SD = 3.86); C M = 5.41 (SD = 2.59); and R M = 7.13 (SD = 2.39).

Predictive validity of the instruments was evaluated using three series of Receiver Operator Characteristic (ROC) analyses. The first set examined the ability of the instruments to predict seclusion, regardless of duration of that stay. A single incident, however, might represent a less serious event, so analyses were also conducted to predict frequent seclusions defined in two ways: above the median (Q1 and Q2 vs. Q3 and Q4) and top quartile (i.e. Q1 vs. Q2, Q3, and Q4). In the second set of analyses, analyses from the first set were repeated with the length of hospital stay taken into consideration in the analyses. The third set evaluated the various instruments’ abilities to identify those whose average seclusion durations (as opposed to number of seclusions) were above the median length or in the upper quartile length. AUC values for each series of analyses were reported by the authors. Overall, ROC analyses indicated that all instruments had small to moderate (AUCs range = .54 to .71) significant predictive validity with respect to frequency of seclusion, but were less strongly predictive of duration of seclusion (AUCs range = .47 to .72).

Project and Scholarly Work


Summary

The current study investigated the predictive validity of the PCL-R, HCR-20, VRAG and LSI-R for frequency and duration of seclusion in a sample of forensic psychiatric patients in Canada. The authors hypothesized that high scores on all four risk-related instruments would be associated with frequency and duration of seclusion, as the practice of secluding patients is often in response to aggressive behavior. The sample used in this study consisted of 130 (116 males and 12 females) adult forensic psychiatric patients found NCRMD who were hospitalized in a medium-secure forensic inpatient unit in Ontario. The participants ranged in age from 22 to 74 years (M = 45.01, SD = 9.87). Risk-related instruments were coded from patients’ institutional files by a registered psychologist trained on the instruments. Incidents of seclusion and description of aggressive incidents were documented by nursing staff and were collected retrospectively from patient files by a pre-doctoral intern in clinical psychology. Seclusion was indexed in terms of total number of seclusions during a period of two years and total time spent in seclusion.
To compare the efficacy of each instrument, the authors rank ordered the predictive accuracy for each of the instruments across the six analyses. Although AUC levels from the various instruments were similar in magnitude, from a content perspective Factor 2 of the PCL-R performed most strongly in the prediction of seclusion of seclusion (mean rank = 1.33). The HCR-20 and LSR-R performed very similarly across the various analyses (mean ranks = 5.16 and 3.83, respectively). The order of rankings for the instruments are as follows: PCL-R Total (2), LSI-R (3), HCR-20 Total (4), H subscale (5), VRAG (6), C subscale (7), Factor 1 (8), VRAG Bin (9) and HCR-20 R (10). Because the efficacy of the measures was compared in a descriptive fashion, whether there were significant differences in predictive utility of the measures was not addressed.

Project and Scholarly Work


Summary

The current study examined the effect of training on the quality of risk assessments using the HCR-20. The research was conducted based out of two forensic psychiatric units. Full HCR-20 training was offered to professionals at the two units. Training involved an initial one day workshop that included an introduction to the tool, case examples and reviews of relevant literature. The training also involved practice cases under the supervision of the trainer and subsequent refresher sessions with the trainer. Following the training, the authors reviewed the patients’ files at the units and rated the quality of the HCR-20 assessments. The reviewer was blind to the training status of the individual who had completed the HCR-20. Each scale of the HCR-20 was rated based on the content and relevant risk information (2 = poor, 3 = good). These reviews included cases that were rated before and after the training. A total of 42 HCR-20 assessments were reviewed for quality.

Following the training it was found that the overall quality of the assessments improved significantly ($t = -6.661$, $p < .001$). As well, the quality of the ratings on each of the subscales also improved significantly. The authors concluded that consideration should be given to providing training for all professionals using structured assessment measures.

Project and Scholarly Work


Abridged abstract (English translation of the study not available):

In this study, the Behavioural Status Index (BEST-Index), an instrument assessing daily living skills and social risk, was investigated and cross validated with the PCL-R and the HCR-20. Participants were 86 German forensic psychiatric patients. All instruments were coded three times over a nine month study period. Sufficient inter rater reliability and good convergent validity of the sub-scales of the BEST-Index was demonstrated in comparison to the HCR-20 and the PCL-R. The authors concluded that clinicians working with the BEST-Index may use it to monitor behavioural change over long treatment periods in mentally ill offenders.

Project and Scholarly Work


Summary

Compared to research focusing on male offenders, the empirical research examining the validity of risk assessment tools in female offenders and female psychiatric inpatients is scarce. The present study investigated the postdictive validity of the HCR-20 and PCL-R in a female forensic psychiatric sample in the Netherlands. Raters scored the measures based on file information and subsequently accessed the participants’ official criminal records. All female patients discharged from two forensic psychiatric hospitals between 1985 and 2001 were included in the study resulting in a sample of 45 former patients. The participants’ mean age when entering the hospital was 28.3 years ($SD = 8.3$). In terms of index offence, 49% had been convicted for homicide, 18% for violent crimes, 22% for arson, and 11%
for property crimes. Outcome data was retrieved from an official criminal database.

The mean HCR-20 total score was 24.8 ($SD = 5.8$), and the mean PCL–R total score was 18.5 ($SD = 6.8$). With respect to the final risk judgments, 24% were classified as low risk, 40% as moderate risk, and 36% as high risk. Interrater reliability was assessed on a subset of 30 cases. Both instruments demonstrated excellent interrater reliability with ICCs of .98 and .97 for the HCR-20 and PCL–R total scores, respectively.

A total of 36% of the participants were reconvicted in the follow-up period with 16% being reconvicted of a violent offence. None of the instruments were postdictive of general or violent recidivism. With respect to general recidivism, the AUCs were .54, .58, .41, and .56, for the total score, H, C and R scales, respectively (all $p < .001$). With regard to violent recidivism, the AUCs were .54, .68, .42, and .56 for the total score, H, C and R scales, respectively (all $p > .13$). Similar, poor postdictive validity was seen for the PCL-R total scores and factor scores (2, 3, and 4 factor variants). For the PCL-R total score, the AUCs were .57 and .60 for violent and general recidivism, respectively. At the item level, only one item was found to be individually postdictive of violent reconviction: H2 (Young age at first violent incident) (AUC = .76).

The results lead the authors to conclude that these risk assessment instruments may not be useful with female forensic psychiatric patients. However, other studies have found the instruments to yield moderate to large effect sizes with female forensic patients (see de Vogel & de Ruiter, 2005) and civil psychiatric patients (see Nicholls et al., 1997, 2001, 2004) and female offenders (see Strand and Belfrage, 2001; Warren et al., 2005).

### Project and Scholarly Work

**Scharin, C. (1999).** *Bedömning av återfallsrisk hos rättspsykiatriskt undersökta personer: En utvärdering av skattningsskalan HCR-20.* Unpublished manuscript.

**Summary**

The Swedish version of the HCR-20 was coded on 49 forensic psychiatric patients. [Sample characteristics unavailable at this time until English translation available]. Proportion of violence in various score categories was calculated for the total HCR-20 score and the H scale alone. Results were as follows: HCR-20 total score from 0-19, 15% violent; total score from 20 to 40, 64% violent. H scale score of 0 to 5 (0% violent), 6 to 10 (31% violent), 11-15 (54% violent), 16 to 20 (80% violent).

**Project and Scholarly Work**


**Summary**

In the UK, the Dangerous and Severe Personality Disorder (DSPD) program was initiated to deal with individuals who have a severe personality disorder and who might pose a risk for future recidivism. However, it is unclear whether those admitted to DSPD programs are different from those admitted to conventional personality disorder (PD) services. In this study, DSPD and PD patients were compared on level of predicted future risk using the HCR-20, previous offending behavior and pre-treatment levels of institutional risk-related behavior. The study sample was comprised of 60 DSPD patients and 44 non-DSPD patients admitted within the same high secure psychiatric hospital in the UK. Mean age of the samples were 33 ($SD = 9.00$) and 34 ($SD = 8.00$), respectively. Both samples were predominantly from a White ethnic group. A majority (93%) of the DSPD sample was diagnosed with a psychopathic disorder, compared to 57% in the PD group. Rates of mental illness and mental impairment were higher in the PD group compared to the DSPD group. DSPD patients received a significantly greater number of convictions after the age of 18, relative to the PD patients, and had been imprisoned on a greater number of occasions. Index offenses of the DSPD patients were more likely to contain violence than those of the PD patients.

HCR-20 assessments were completed on admission to either the DSPD or PD unit of the hospital as part of routine clinical procedure. DSPD patients’ data on offending history was collected during admission and PD patients offending history was collected from review of hospital records. The initial 6 – 12 months following admission to either of the unit consisted of patient assessment and case formulation, and was therefore referred to by the authors as the ‘pre-treatment period.’ Pre-treatment institutional risk-related behavior was collected from hospital incident reports over the 12-month period. These incidents were coded into the following subtypes: physical interpersonal aggression, verbal aggression, and total institutional risk-related behavior that combined the two categories above.
Inter-reliability computed on a subsample of 20 randomly chosen incidents was good (ICC = 0.93). Within the first 12 months of admission, DSPD patients engaged more frequently in physical interpersonal aggression, verbal aggression and all institutional risk-related behavior relative to PD patients.

Ten patients within the sample (7 PD patients and 3 DSPD patients) had not been assessed using the HCR-20 within two years of their admission date and were excluded from any analyses using these assessments. Relative to PD patients, DSPD patients obtained significantly higher scores on the HCR-20 scale, and on the clinical and risk subscales, but scores on the historical scale did not significantly differ between groups. With respect to the PD group, median scores were 24.0 (IQR = 8.0), 17.0 (IQR = 4.0), 4.0 (IQR = 4.0) and 2.0 (IQR = 2.0) on the total, historical, clinical and risk management items of the HCR-20. With respect to the DPSD group, median scores were 28.0 (IQR = 8.5), 16.0 (IQR = 2.0), 6.0 (IQR = 2.0) and 7.0 (IQR = 2.0) on the total, historical, clinical and risk management items of the HCR-20.

There was evidence to suggest that the time taken to complete HCR-20 assessment after admission of the PD sample was significantly longer (M = 239 days, SD = 167 days) than for the DSPD sample (M = 138 days, SD = 65 days). Therefore, the PD sample could have obtained lower scores on the HCR-20 relative to the DSPD sample because their assessments were conducted after a greater inpatient stay. Statistically controlling for time until completion of the HCR-20 assessment, the authors found that higher scores on the HCR-20 scale, clinical and risk subscales were significantly associated with admission to a DSPD facility.

Spearman’s rho correlation analysis revealed that the total incidents of risk-related behavior during the initial 12 months of admission were significantly correlated with HCR-20 total scores ($r_s = 0.42, p < .001$), clinical subscale scores ($r_s = 0.47, p < .001$), and risk subscale scores ($r_s = 0.40, p < .001$). Likewise, these scales were associated with incidents of interpersonal physical aggression ($r_s = 0.28, p < .01$; $r_s = 0.35, p < .001$; $r_s = 0.21, p < .05$, respectively) and verbal aggression ($r_s = 0.44, p < .001$; $r_s = 0.44, p < .001$; $r_s = 0.34, p < .001$, respectively). The HCR-20 historical subscale scores also correlated significantly with incidents of verbal aggression ($r_s = 0.24, p < .05$). However, the historical subscale did not correlate significantly with total incidents of risk-related behaviour ($r_s = 0.12$), or interpersonal physical aggression ($r_s = 0.13$).

**Project and Scholarly Work**


**Summary**

To investigate whether care of forensic psychiatric patients in Auckland, New Zealand was more related to treatment for their illness or punishment for their offending, this study examined whether clinical progress (operationalized as access to unsupervised leave) was associated more with clinical factors or with criminological factors (e.g., time served proportional to the severity of offending). A non-experimental cross-sectional study design was used with this sample that comprised all mentally disordered offenders in the Auckland region under forensic care (96 patients, 74 of whom were inpatients).

Participants’ mean age was 35.7 years ($SD = 9.23$, range 18-62). Most participants were men ($n = 88; 91.7%$). There were no significant differences in mean age or gender between participants who were or were not granted unsupervised leave. More than half (52.1%) of the sample was New Zealand Maori (36.4% European; 11.5% Pacific Islanders). Maori (56.0%) and Pacific Islanders (81.8%) were significantly more likely to be restricted than Europeans (34.2%; $\chi^2 = 8.59, p = .01$).

To quantify severity of offending, a Crown prosecutor assisted in calculating a theoretical custodial sentence and date of parole using information from an offence summary or police summary of facts. A treating psychiatrist made DSM-IV diagnoses. Severity of mental disorder was assessed with the Health of the Nation Outcome Scale (HoNOS; Wing et al., 1998). General functioning was measured with the Life Skills Profile (LSP-39; Rosen, Hadzi-Pavlovic, & Parker, 1989), which comprises five categories: self-care, non-turbulence, social-contact, communication, and responsivity. Clinicians who gathered the data were not blind to patients’ leave status. Inter-rater reliability was assessed before the data were collected (values not reported). The psychopathy item (H7) was omitted when scoring the HCR-20.

Eighty-four participants were diagnosed with psychotic-spectrum disorders. Of five participants who did not have a diagnosis on Axis I, three had personality disorder diagnoses and two had mild mental retardation. There were no significant differences on Axis I diagnoses between the two leave groups ($\chi^2 = 1.87, p = .76$).

There was not a victim in 16.7% of cases. Strangers (28.1%), acquaintances (28.1%), and family members (27.1%) were victimized in similar proportions. Victim type
did not differentiate the two leave groups ($\chi^2 = 4.38, p = .22$).

Inspection of HCR-20 scores indicated that scores on the total measure and on the Clinical and Risk Management scale scores, but not on the Historical scale, differed significantly between participants who were or were not granted access to unsupervised leave. Mean HCR-20 scores were as follows: Total (detained = 25.23, SD = 7.11; released = 18.26; SD = 5.06; $p < .01$); Historical (detained = 13.67, SD = 3.30; released = 12.94; SD = 3.51; $p = .30$); Clinical (detained = 5.56, SD = 2.81; released = 2.83; SD = 2.28; $p < .01$); Risk Management (detained = 5.88, SD = 2.72; released = 2.49; SD = 1.88; $p < .01$). ROC analyses were consistent with these results and revealed that historical risk factors were not discriminatory of leave status (AUC = .56, SE = .06, $p = .31$, 95% CI: .45-.68). The Total (AUC = .77, SE = .05, $p = .00$, 95% CI: .68-.87), Clinical (AUC = .76, SE = .05, $p = .00$, 95% CI: .67-.86), and Risk Management (AUC = .85, SE = .04, $p = .00$, 95% CI: .78-.92) scales were predictive of leave status.

The two leave groups did not differ significantly in terms of severity of offending ($t = -0.3, p = .97$), time served ($t = -0.65, p = .52$), or time served relative to offending severity ($t = -0.63, p = .53$). Offence type significantly differentiated the groups ($\chi^2 = 13.63, p < .05$), with sex offenders being significantly more likely to be detained compared to other types of offenders.

Binary logistic regression was used to examine the ability of demographic (i.e., age, gender, and ethnicity), clinical (i.e., scores on HoNOS, HCR-20, and LSP-39), and criminological (i.e., legal status, type of offence, offence severity, time served, proportion of time served relative to offence severity) variables to predict leave status. None of the criminological factors reached statistical significance and of the demographic variables, only ethnicity reached significance ($R^2 = .12, p = .01$). Clinical factors – especially those assessed by the HCR-20 Risk Management scale - were most predictive. Values for the HCR-20 indices were as follows: Total ($R^2 = .32, p < .01$); Historical ($R^2 = .02, p = .27$); Clinical ($R^2 = .29, p < .01$); Risk Management ($R^2 = .45, p < .01$).

In summary, results indicated that dynamic clinical and risk assessment variables had improved among participants granted release, whereas static and criminological variables were not significantly different between the two leave groups.

**Summary**

The goals of the present study were: (1) to identify if patients in a low security forensic psychiatry unit in Scotland who were admitted from a high security hospital were inpatients for a longer period of time than patients admitted from other sources (e.g., prison, other hospitals, and police custody) and (2) to examine whether any individual factor could predict a length of inpatient stay of more than 2 years (i.e., the length of time recommended by a local policy report). Using a retrospective cohort design, all patients ($n = 17$) transferred from a special security unit in Scotland ($n = 16$) and in England ($n = 1$) between 1990 and 2002 were compared to a control group that comprised 17 consecutively admitted patients from any other referral source. File information was used to code demographic details and offending history. The last clinical diagnosis recorded on the multidisciplinary team review was coded for the present study. The HCR-20 was completed, with the psychopathy item (H7) omitted, for all participants with four exceptions in the control group - for two participants, only the Historical scale was completed because they had died (one from natural causes and one from suicide), and for another two participants, insufficient documentation prevented scoring of all HCR-20 indices.

There was no difference in age between participants transferred from the special security unit ($M = 40.5$ years) and participants in the control group ($M = 36.8$ years). The average HCR-20 total score was significantly higher for the special hospital group ($M = 27.5$) than for the control group ($M = 20.7$), $p < .005$. There was a significant difference in diagnosis between the two groups, $\chi^2 = 7.7, df = 4, p = .01$. Ten special hospital patients were diagnosed with schizophrenia compared to 4 control patients. The type of index offence also differed significantly between the two groups, $\chi^2 = 9.6, df = 6, p = \text{invalid due to small numbers},$ with violent index offences being more common in the special hospital group.

The outcome of inpatient stays was significantly different between the two groups, ($\chi^2 = 16.6, df = 5, p < .005$). Among the 17 patients transferred from special hospitals, 11 remained inpatients in forensic service, compared to only one participant in the control group. The mean length of stay for the special hospital group was 2.41 years ($SD = 2.9$ years, range = 2 weeks-11 years). The mean length of stay for the control group was 0.55 years ($SD = 1.4$ years, range = 1 day-6 years).
A regression analysis to predict length of stay was completed with the following variables: HCR-20, age, age at first symptoms, diagnosis, index offence, and previous offences. Although the overall model was significant (adjusted $R^2 = 0.04$, $F = 1.11, p < .05$, 95% CI: 0.18-10.64), no single factor was significant in predicting length of stay.

In summary, patients transferred from special hospitals to the low security forensic unit were more likely to have a diagnosis of schizophrenia, a more serious index offence, a lengthier criminal history involving violence, and a higher HCR-20 score compared to patients admitted from other sources. They also were more likely to remain as inpatients in forensic service.

**Project and Scholarly Work**


**Summary**

This study examined the predictive validity of the VRAG and HCR-20 in a sample of mentally disordered patients from a black ethnic minority using a pseudo-prospective case note analysis design. The final study sample consisted of 1,016 mentally disordered offenders (834 Caucasian and 249 Black patients) discharged from medium-secure psychiatric facilities in the UK between December 1992 and September 2001. The black and white groups were well matched in terms of gender (males: white = 83%, black = 86%) and average age (white = 32.0, black = 31.2). The prevalence of mental illness was smaller in white than black patients (65.7% vs. 87.6%), while there was a greater prevalence of personality disorder (27.2% vs. 9.2%) and intellectual disability in the white participants (14.1% vs. 5.6%).

The HCR-20 and VRAG were coded using file-based information by four psychologists blind to outcome following discharge. On a test sample of 19 cases ICC values were 0.80 and 0.95 for the HCR-20 and the VRAG, respectively. The main dependent variable in the study was occurrence of an offense over a two-year period following discharge from secure psychiatric services and was obtained from the UK Ministry of Justice Offenders Index. Offenses were grouped as whether they were violent or any offense (which also included violent offenses). Violent offenses included all offenses classified as violence against the person by the Home Office and kidnap, criminal damage endangering life, Robbény, rape, and indecent assault. Time to offense was also calculated and was operationalized as the difference between the discharge date and the time of reconviction.

Mean scores and AUC values for violent offending were reported. For the sample as a whole, mean scores were 4.73 ($SD = 10.25$), 18.5 ($SD = 6.5$), 11.3 ($SD = 3.7$), 3.3 ($SD = 2.5$), 3.8 ($SD = 2.6$), and 4.73 ($SD = 10.25$) on the VRAG, and HCR-20 total, historical, clinical, and risk management scales, respectively. Black participants appeared to have lower risk scores on both the VRAG and the HCR-20 than the white participants (Cohen’s $d$ = 0.20 – 0.38), however these differences were not statistically significant. With regards to the white participants, mean scores were 5.55 ($SD = 10.51$), 19.0 ($SD = 6.9$), 11.5 ($SD = 3.8$), 3.5 ($SD = 2.5$) and 3.8 ($SD = 2.6$) on the VRAG and HCR-20 total, historical, clinical, and risk management scales, respectively. With regards to the black participants, mean scores were 2.44 ($SD = 7.98$), 16.9 ($SD = 6.0$), 10.7 ($SD = 3.5$), 3.0 ($SD = 2.5$) and 3.0 ($SD = 2.2$) on the VRAG and HCR-20 total, history, clinical, and risk management scales, respectively.

AUC values for the VRAG and the HCR-20 were mainly good for both white and black participants. Furthermore, differences in AUC values between the two groups were small and not statistically significant. For black participants, AUC values were .74, .66, .68, .53, and .62 on the VRAG and HCR-20 total, historical, clinical, and risk management scales, respectively. For white participants, AUC values were .79, .72, .71, .54, and .69 on the VRAG and HCR-20 total, historical, clinical, and risk management scales, respectively. For the sample as a whole, AUC values were .76, .71, .70, .54, and .69 on the VRAG and HCR-20 total, historical, clinical, and risk management scales, respectively.

A survival analysis was conducted to determine racial differences in time to violence. A Kaplan-Meier Log Rank test showed that there were no significant differences in overall survival rates (white = 3278 days, black = 2932 days). There were no significant differences between groups in terms of conviction for any offense (19% of white offender’s vs. 16% of black offenders were reconvicted) or for a violent offense (12% of white offenders vs. 11% of black offenders were reconvicted).

The authors concluded that the HCR-20 and VRAG were accurate for predicting violent reconvictions in a sample of black patients discharged from medium psychiatric units in the UK. They also noted that differences in risk scores may reflect different rates of mental illness and personality disorder between the two groups. Future research should investigate whether the HCR-20 and the VRAG are consistency accurate for predicting violence in black
minority groups, as well as whether these measures are accurate across other ethnic minority groups.

Project and Scholarly Work


Summary

This study examined the reliability of the Brazilian version of the HCR-20. Participants were randomly selected from a single unit at a forensic psychiatric hospital in Brazil. Two psychiatrists assessed 30 males using the HCR-20. The results indicate excellent interrater reliability. The ICCs were .97 for the H scale, .94 for the C scale, .96 for the R scale and .96 for the total score. Looking at the item level, the ICCs ranged from good to excellent with a mean ICC = .97. Item level ICCs ranged from .60 to .99, with 18 of the items having an ICC of .84 or greater. Overall, the reliability of the Brazilian version of the HCR-20 was supported.

Project and Scholarly Works


Summary

This study investigated the H scale of the HCR-20, the VRAG and the PCL-R in 169 forensic patients with schizophrenia detained in a state hospital between 1992 and 1993. Due to exclusions (death, unavailable or poor quality of records) the final sample was 140 participants.

The outcome of interest was ‘any’ incident defined as any aggressive incident involving physical contact with a victim, any sexual incident, or any episode of physical aggression towards property. Next, a ‘serious’ incident was defined as any aggressive incident resulting in death or injury requiring hospital treatment, any sexual incident involving contact with the victim, or any fire setting. In addition, an offence was defined as any conviction (including non-violent offences) and ‘violent’ offence was defined as any conviction for assault, serious assault, fire-setting/raising or a contact sexual offence. Criminal records were only available for 135 patients, the Krawieka Rating Scale was completed in 132 cases and all other analyses were of 140 cases. The sample was mostly male (90%) and the mean age was 35.4 years (range 19-63 years). Sixty-five percent of participants had a comorbid diagnosis with the most common being antisocial personality disorder, alcohol dependence and drug dependence. Alcohol abuse was prevalent in 58% of participants and 54% had drug abuse. The majority of participants had been previously convicted for a crime (26.4% homicide and 19.3% sexual offence). The average length of stay (prior to January 1994) was 6.1 years (range .6-26.3 years).

Interrater reliability was assessed in 9 cases, four by one researcher and five by another. Correlation coefficients revealed good interrater reliability for all three measures. For the H-10 (H subscale of the HCR-20) the rs were .843 (p = .079) and .884 (p < .01). For the VRAG the rs were .946 (p < .01) and .703 (p = .052). For the PCL-R the rs were .827 (p = .084) and .714 (p = .071).

Within the sample, 76.4% had at least one incident and 27.9% had at least one serious incident. Most incidents were against staff or other patients in the hospital. Among participants living in the community, 14.8% committed a new offence and 5.1% committed a violent offence. Of the 107 patients who were discharged, 21 were readmitted due to re-offending or violence in less secure in-patient settings. The mean of the H-10 was 13.4 (SD = 3.4). The mean of the VRAG was 2.3 (SD = 10.6) and the mean for the PCL-R was 14.3 (SD = 7.1). The H-10 mean scores were not significantly different for those who did commit any incident or any serious incident compared to those who did not. The H-10 mean scores were significantly different for any conviction and any violent reconviction. The same pattern emerged for the VRAG and the PCL-R. The H-10 did not significantly discriminate between those who did or did not leave the high security hospital or between those who were and were not readmitted after discharge. It did significantly differentiate between those who reached the community after leaving the hospital. In the ROC analysis, neither the H-10, the PCL-R nor the VRAG significantly predicted ‘any’ or ‘serious incidents.’ All three did significantly predict ‘any’ (H-10 AUC = .76, VRAG AUC = .76, PCL-R AUC = .73) and ‘violent offences’ (H-10 AUC = .79; VRAG AUC = .80, PCL-R AUC = .83). The same pattern emerged for those patients who actually got to the community (n = 54; H-10 AUC = .77-.80; VRAG AUC = .76-.77; PCL-R AUC = .78-.84). None of the instruments predicted frequency of incidents and serious incidents.

Summary

This prospective study examined the predictive validity of the HCR-20 when applied to clinical practice among 109 male mentally disordered offenders in a high secure forensic hospital in Scotland. The HCR-20 was rated by clinicians as a part of standard practice in the hospital. The mean age of the sample was 38.6 years (SD = 10.7). A majority of the sample had a psychotic illness (92.7%), dissocial personality disorder (61.5%) and was single (80.7%) at the time of the assessment. Research participants were followed up for a mean of 31 months (SD = 8.3) post recruitment. During this time, 60.6% of participants were discharged from the hospital: 81.8% to medium secure facilities, 18.2% to low secure settings and 16.7% to the community. All participants were followed up across less secure settings if discharged during the study period. The main outcome variables were any violent incident recorded in a patient’s hospital files, official reconviction data, or hospital incident reporting systems. In total, 234 incidents were committed by 42.2% of participants. Of all incidents, 79.1% involved violence against others, 13.7% were sexual incidents, and 7.2% involved violence against property. Of the patients discharged to the community, 7.6% were reconvicted and 25.8% were readmitted to a psychiatric hospital.

The HCR-20 profile for the sample was compared to data from participants who declined to participate in the research study (n = 36). With regards to the study sample, mean scores were 25.22 (SD = 5.1), 15.30 (SD = 2.70), 5.22 (SD = 2.4) and 4.70 (SD = 2.4) on the total, H, C, and R scales of the HCR-20, respectively. For non-respondents, mean scores were 27.46 (SD = 5.9), 15.15 (SD = 3.7), 6.33 (SD = 2.3) and 5.97 (SD = 2.90) on the total, H, C, and R scales, respectively. Non-respondents were associated with a higher overall risk of future violence than participants (U = 1504.5, p < .05). In particular, non-respondents were rated as clinically more unwell (U = 1441.0, p < .05) and requiring more intensive risk management (U = 1419.5, p < .05) than those who agreed to take part in the study.

The efficacy of the HCR-20 was assessed using Receiver Operating Characteristic (ROC) analysis. AUC values were reported for each scale of the HCR-20 and the following four outcomes: all incidents, minor incidents, serious incidents, and any conviction. For all incidents of violence and all minor incidents of violence the HCR-20 was a poor predictor in this study. With regards to all incidents, AUC values were .50, .54, .55 and .51 for the total, H, C, and R scales, respectively. With regards to all minor incidents, AUC values were .54, .56, .52 and .53 for the total, H, C, and R scales, respectively. For serious incidents, while the AUCs on all subscales were relatively high (AUC = .68, .79 and .75 for H, C, and R scales, respectively), only the HCR-20 total score was a significant predictor (AUC = .86, p < .05). For any conviction the AUC values were .60, .56, .61 and .55 on the total, H, C, and R scales, respectively.

Cox regression analyses using the HCR-20 scales, with time at risk calculated in months as the duration between the baseline HCR-20 and the first incident, were conducted. Time dependent covariates were age at baseline and length of follow-up. The mean time between implementation of the HCR-20 and the first incident was 10.8 months (SD = 8.9). The Cox regression model was statistically non-significant for all and minor incidents. None of the predictors (HCR-20 total score, H score, C score, R score, age and follow-up time) predicted incidents. The same findings were obtained when the follow-up period was categorized into specific intervals (i.e., 1 month, 12 – 24 months, 24 – 26 months, and > 36 months). No analysis was conducted for serious incidents due to the small number (n = 3).

The results imply that the implemented HCR-20s did not predict future violence regardless of setting (community vs. inpatient) or time (short vs. long term) except for serious incidents. After examining findings obtained from an earlier study with a similar cohort, the authors concluded that results may indicate that the implemented HCR-20s informed risk management through systematic tailoring and treatment plans. Alternatively, the completion of violence risk assessment by clinicians rather than researchers may have affected the quality of completed assessments. It should be pointed out that these findings may be study-specific, given that Guy’s (2008) meta-analysis showed that the HCR-20’s association with violence was not affected by whether it was completed for research or clinical purposes.

See Also


Project and Scholarly Work


Summary

This dissertation aimed to establish normative standards for a new measure of violence risk assessment comprised solely of dynamic variables, the Violence Clinical Risk Indicator (VCRI), in a sample of forensic psychiatric inpatients. As part of this study, the VCRI was compared the PCL-R and HCR-20. Fifty-four forensic psychiatric patients in Oklahoma completed the study. All were adjudicated as Not Guilty by Reason of Insanity (NGRI). The mean age of the participants was 43.13 years (SD = 11.05). Participants were primarily male (55.5%), Caucasian (55%), and single (59%) at the time of the study. The main primary diagnosis in the sample was schizophrenia (40.74%). Approximately 31% of the sample had a co-occurring substance abuse or dependence diagnosis, while 22.2% of the participants had a comorbid diagnosis of antisocial personality disorder.

Participants were assessed at pre-test with the VCRI, PCL-R, and HCR-20. For some participants, pretest scores for the PCL-R and HCR-20 were based on previous administrations in the hospital.

Internal consistency values, as indicated by Cronbach’s alpha, were 0.85, 0.84 and 0.79 for the VCRI, PCL-R, and HCR-20, respectively. Violent and non-violent behaviors were recorded weekly from progress notes and critical incident reports over a 6-month period. To avoid overestimating aggression behavior, violence was divided into two categories. Violence in category 1 consisted of actual physical violence, physical assault, sexual assault, destruction of property, violent acts with the use of a weapon, and violent threats with a weapon (if a weapon is in hand or is present). Violence coded in category 2 included aggressive or violent acts that do not result in an injury. The mean number of Category 1 violent incidents in the sample was .81 (SD = 1.67), the mean number for Category 2 violent incidents was 2.26 (SD = 7.08), and the mean number for Total violent behavior was 3.07 (SD = 8.09). Thirty-seven of the participants did not demonstrate Category 1 violent act, 28 participants did not demonstrate Category 2 behavior and 28 of the participants did not demonstrate any Total violent behavior.

Mean scores were 37.37 (SD = 11.91), 15.78 (SD = 8.04) and 19.65 (SD = 6.46) on the VCRI, PCL-R and HCR-20, respectively. Mean scores on the HCR-20 subscales were 12.37 (SD = 3.72), 4.91 (SD = 2.62) and 2.35 (SD = 2.18) for the historical, clinical and risk management scales, respectively. Total score on the VCRI was significantly correlated with both PCL-R total score (r = .38, p < .01) and HCR-20 total score (r = .57, p < .01).

A hierarchical regression analysis was conducted to examine whether the VCRI predicted cumulative violent behavior more accurately than the PCL-R or the HCR-20. Since the total score from the VCRI had a stronger relationship with total violent behavior, it was the first variable entered into the Stepwise equation. The total score from the HCR-20 also produced significant results, (AUC = .67, p < .05). Total score on HCR-20 also produced significant results, (AUC = .67, p < .05), while the PCL-R failed to significantly predict cumulative violence, (AUC = .57, p > .05). Total score on VCRI significantly predicted Category 1 violent behavior (AUC = .70, p < .05). Total score on HCR-20, (AUC = .68, p < .05), as well as its Historical scale, (AUC = .68, p < .05), also produced significant results for Category 1 violent behavior. The total score from the PCL-R failed to produce significant results, (AUC = .54, p > .05); however, Factor 2 from the PCL-R significantly predicted Category 1 violent behavior, (AUC = .71, p < .05). None of the total scores from the measures significantly predicted Category 2 violent behavior. Only the Risk Management scale from the HCR-20 produced significant results, (AUC = .67, p < .05).

Although these findings are highly supportive of the utility of the VCRI, the author notes some limitations to the findings. First, because HCR-20 scores were based on previous administrations for some of the participants predictive validity of these scores may have been affected. Second, violent behaviors may have been unrecorded due to inaccurate records or due to violent behaviors that occur without being noticed by hospital staff.

Project and Scholarly Work

Summary

This study examined the utility of the HCR-20 and PCL-R in predicting institutional violence in a secure psychiatric facility in Norway. Participants were 44 men (mean age = 31.8 years) and 7 women (mean age = 25.7 years) admitted over a 10 year period. Most participants had a primary diagnosis of schizophrenia (41%) or other psychosis (47%). The remaining 12% had a primary diagnosis of a personality disorder.

Mean HCR-20 scores were: Total (23.5, SD = 6.8); H-scale (13.8, SD = 4.3); C-scale (5.9, SD = 1.9); R-scale (3.9, SD = 2.0). Mean PCL-R scores were: Total (19.4, SD = 8.8); Factor 1 (7.6, SD = 3.5); Factor 2 (9.1, SD = 5.2).

The Staff Observation Aggression Scale was used to code aggressive episodes. Correlations were computed between the frequency of aggression (i.e., total episodes divided by patient days), severity of episode, and occurrence of physical aggression. The HCR-20 total score, C-scale, and R-scale (risk in institution) were correlated significantly with frequency (rs between .36 and .40), and there was a trend toward significance for the H-scale (r = .28, p = .06). The HCR-20 total score, H-scale, and R-scale were correlated significantly with severity (rs between .38 and .44). The HCR-20 was not correlated significantly with physical aggression (r = .29, p = .08); nor was the PCL-R (PCL-R F2, r = .28, p = .08). The only significant correlations obtained with the PCL-R were between frequency and total score and Factor 2 and between severity and total score. Frequency and proportion of physical aggression were significantly higher among women than men.

AUC values for the prediction of frequency were: HCR-20 total (.76); H-scale (.67); C-scale (.82); Risk (.70); PCL-R Factor 1 (.64); PCL-R Factor 2 (.77). AUC values for the prediction of most severe episode were: HCR-20 total (.82); H-scale (.77); C-scale (.73); Risk (.76); PCL-R total (.73); PCL-R Factor 1 (.65); PCL-R Factor 2 (.71).

Project and Scholarly Work


Abridged Abstract (English translation of the study not available):

Background: Instruments for assessing the risk of violent behaviour have proved their worth in the field of forensic psychiatry. However, it is not certain whether such instruments provide an accurate assessment of recidivism when used with psychiatric patients who have an intellectual disability. Aim: To find out whether these instruments – or which of these instruments – can provide a reliable assessment of the risk of violent sexual recidivism when used with forensic psychiatric patients who have an intellectual disability. Method: We searched PubMed, Psycinfo and Google Scholar for studies in English or Dutch, published between 1980 and 2010, concerning the reliability of instruments for assessing the risk of recidivism in forensic psychiatric patients with an intellectual disability. Results: The studies show that several of the risk assessment instruments currently in use provide reliable predictions concerning patients with an intellectual disability. Conclusions: The HCR-20 and HKT-30 are recommended for assessing the risk of violent recidivism in patients with an intellectual disability. The Static-99 is the preferred instrument for assessing the risk of sexual recidivism. Further research is needed for assessing to what extent SVR-20, Armidilo, Dros and PCL-R are applicable to forensic psychiatric patients in the Netherlands who have an intellectual disability.

Project and Scholarly Works


Summary

This is a chart review study of 250 persons referred from court to a maximum security forensic institute over the course of five years for the purpose of assessment of criminal responsibility. The focus of this summary is the 125 persons who were found Not Criminally Responsible on Account of Mental Disorder (NCRMD) for their offences. The sample (M age = 34.98; SD = 10.67) was primarily male (82.4%), Caucasian (77.4%), single (88.6%), unemployed (76.4%), and many patients had less than grade 11 education (40.2%). Most patients had
committed a violent index offence (77.6%), and most had a primary diagnosis of a psychotic disorder at assessment (66.9%), followed by mood disorder (21.0%).

The purpose of the study was to evaluate which factors predicted (1) verdicts of NCRMD (insanity acquittal) versus guilty, (2) length of confinement and days in the system. The HCR-20 was use for the latter, as a predictor of days in the system and days confined. Also included in such analyses were a variety of criminological, psychiatric, demographic variables, and the Psychopathy Checklist: Screening Version. Hierarchical Cox Proportional Hazards Regression was use as the method of prediction, with time in the system as the dependent measure. After all blocks and variables were entered, the HCR-20 was the only significant predictor, with an $e^b = .898$ (Odds = 2.45). In particular, the R Scale was the strongest of the three scales. Using a somewhat more liberal approach with a backward elimination entry procedure, one other variable in addition to the HCR-20 entered the equation (offence severity). For a slightly different dependent measure (days until first release), several variables entered the model (using backward elimination): level of violence, number of remand charges, homicidal at offence, age at first mental health contact, PCL:SV, and HCR-20.

Months confined, in the system, and until first release were calculated as a function of low, moderate, and high scores on the HCR-20 (by dividing the total scores into thirds). Months in the system, confined, and until first release, for the LOW group were 32.82, 9.22, and 7.25, respectively. For the MODERATE group, results were 38.68, 18.56, and 13.93, respectively. For the HIGH group, results were 45.47, 40.23, and 30.92, respectively. These findings provide support for the concurrent validity of the HCR-20. The factors it predicted are related to legal concepts of risk and threat.


Summary

This study is based on the same data set as Vincent (1998), infra, but addressed independent issues. The reader is referred to the annotation of Vincent (1998) for a description of the general methodological factors.

This research assessed the predictive ability of the HCR-20, VRAG, and PCL:SV in terms of inpatient violence of 125 forensic psychiatric patients. The authors carried separate analyses for pre-disposition and post-disposition time periods (i.e., pre- and post Review Board hearing). Violence was categorized as verbal, physical, and “any,” and was coded from detailed files. Analyses included univariate Pearson r correlations, ROC analyses, and hierarchical logistic regression analyses.

Pre-disposition violence. For the HCR-20, Pearson r values for verbal, physical, and any violence were as follows: .39, .36, and .46. These generally were higher than for the VRAG (.22, .07, and .21) or the PCL:SV (.25, .26, and .32). The AUCs for the HCR-20 were .72, .72, and .77, and again were generally higher than for the VRAG (.62, .66, and .69) or PCL:SV (.65, .54, and .62). Hierarchical logistic regression showed that the PCL:SV predicted any and physical inpatient violence when entered as alone in Block 1, the VRAG did not add to this on Block 2, and, on Block 3, only the HCR-20 predicted violence (the PCL:SV was no longer significant, nor was the VRAG). Results were not reported for verbal violence.

Post-disposition violence. For the HCR-20, Pearson r values for verbal, physical, and any violence were as follows: .31, .31, and .36. These generally were higher than for the VRAG (.20, .08, and .23) or the PCL:SV (.20, .14, and .16). The AUCs for the HCR-20 were .68, .69, and .71, and again were generally higher than for the VRAG (.62, .55, and .63) or PCL:SV (.60, .58, and .59). Hierarchical logistic regression showed only the HCR-20 predicted any and physical violence (the PCL:SV and VRAG were not significant in any Block). Results were not reported for verbal violence.


Summary

This is a chart review study of 172 persons either found unfit to stand trial ($n = 50$) or not criminally responsible on account of mental disorder (NCRMD; $n = 122$). The sample ($M \text{ age } = 37$) was primarily male (83.14%), Caucasian (80.23%), single (88.37%), and unemployed (79.07%). Most patients had committed a violent index offence (75.58%), and most had a primary diagnosis of a psychotic disorder at assessment (60.47%), followed by bipolar disorder (16.28%).
The purpose of the study was to evaluate which factors predicted criminal review board release decisions (discharge versus custodial detention). A variety of mental health, criminological, and demographic characteristics were used as predictor measures along with the HCR-20 and PCL:SV. Hierarchical logistic regression was used as the method of prediction, with release decision as the dependent measure. For the first review board hearing (patients have regular hearings until released), the H Scale, C Scale, and R Scale were entered in separate blocks. Each was a significant predictor (lower scores relating to discharge), with C and R adding incrementally to H. R was most strongly related to discharge decisions. Additional analyses were carried out to predict subsequent discharge/custody decisions. Three Clinical subscale items (Negative Attitudes, Lack of Insight, and Impulsivity), and one Risk Management scale item (Noncompliance with Remediation Attempts) predicted discharge.

These results suggest that at the first hearing, the Risk Management items were most important for discharge decisions, although the Clinical and Historical items also were predictive. At subsequent hearings, change in mental status (Clinical Scale items) emerged as the more important predictor. Results provide support for the concurrent validity of the HCR-20. Release decisions legally require the Review Boards to take into account the threat posed by the individual, the need to reintegrate the accused into society, and the mental condition of the accused.

Project and Scholarly Work


Summary

Vincent et al. (2001) investigated the correspondence between HCR-20 (Version 1) C and R scale ratings made by psychiatrists from file + interview, and version 2 ratings made by researchers from file alone. There was a significant difference between clinicians and researchers on the C scale, but not on the R scale. The difference, though significant, was small (Cohen’s $d = .31$). The association between rater groups was fairly high ($ICC = .58$ for C scale; $ICC = .70$ for R scale). There were few (0-4 per item) “0-2” coding disagreements on individual items. Vincent et al. (2001) concluded that the file-based ratings were sufficiently reliable for research purposes.

Project and Scholarly Work


Summary

The authors noted that pre-admission forensic nursing assessment does not have an empirical evidence base despite the fact that such assessments are completed routinely and at substantial financial cost. The purpose of this study was to assess retrospectively the quality of pre-admission risk assessments completed by nurses at one forensic psychiatric clinic in Wales through comparison with the HCR-20, Version 1.

Eighty-five consecutive cases referred for forensic nurse assessment over a 51-month period were coded as to whether sufficient information had been collected during the risk assessment to allow scoring of the HCR-20 items. HCR-20 item 7 (PCL-R score) was omitted because the researchers did not have formal training in administration of the PCL-R. In roughly 89% of cases, there was sufficient information to rate the 5 items comprising the C-scale. Pertaining to the H-scale, 7 of the 9 items could be rated in over 80% of the cases. Presence of a personality disorder and employment history could be rated in 71% and 77% of the cases, respectively. Four of the five R-scale items could be coded for over 85% of the cases. The Risk management item that assesses the feasibility of plans was able to be rated for less than 80% of the cases.

Mean scores were: HCR-20 Total $23.6$ (3.6); H-scale $12.5$ (2.8); C-scale $5.5$ (3.4); R-scale $5.5$ (3.1). In the discussion section the authors concluded that information traditionally collected by forensic nurses in the course of a risk assessment was appropriate, but noted the importance of research-based practice.

Project Description

Webster, C. D., Hart, S. D., & Eaves, D. Prospective study of the HCR-20 in a forensic psychiatric setting.
Summary

This is a prospective study. The HCR-20 was coded on 175 consecutive persons who were coming before a Criminal Review Board for release from dispositions of Not Criminally Responsible an Account of Mental Disorder (NCRMD). The PCL-R was coded with the use of interviews by trained assistants. Psychiatrists who were responsible for providing the Board with a release assessment completed the Brief Psychiatric Rating Scale (BPRS), as well as the Clinical and Risk Management scales of the HCR-20, as part of their assessments. The H scale was coded by assistants on the basis of file and interview information.

The mean age at admission was 33.0 (SD = 9.6). The sample was primarily male (n = 133, 88.7%). The vast majority of participants were unemployed at admission (n = 139, 92.7%). Schizophrenia was the predominant admission Axis I diagnosis (n = 96, 64%). Forty-one patients (27.4%) of the patients received an admission diagnosis of personality disorder. Most patients had been hospitalized in the past (n = 132, 88%), and the majority had previous charges for violent offences (n = 90, 60%). Finally, most patients had a violent index offence (n = 129, 86%). Violence was measured in the hospital with the Overt Aggression Scale, and in the community with arrests records and re-admission to the forensic institute.

Scholarly Works


Summary

The HCR-20 violence risk assessment scheme was coded on 175 consecutive insanity acquittees appearing before a criminal Review Board. The purpose of the study was to provide data on the descriptive, normative, and reliability characteristics of the HCR-20, and on its relationship to conceptually-related concurrent measures and indexes. The alpha co-efficients for the HCR-20 Total, H scale, C scale, and R scale scores, respectively, were .78, .69, .77, and .77. Other indexes also supported the structural reliability of the HCR-20 (i.e., MIC; CITC). For the H Scale, interrater reliability was good (ICC1 = .81; ICC2 = .90). Interrater reliability was not available for the other HCR-20 scales. Test-retest analyses showed that the C and R scales changed (declined) across repeated assessments, as they are expected to.

The HCR-20 was related strongly to the PCL-R, correlating at .60. The H Scale was most strongly related (.76 with PCL-R Total), while the C and R Scales were related with small effect sizes (rs = .18 and .16, respectively). The H Scale was more strongly correlated with Factor 2 of the PCL-R, while the C and R Scales were more strongly correlated with Factor 1. The HCR-20 and its scales were related to psychopathology (Brief Psychiatric Rating Scale; various factors), in a conceptually meaningful way. Generally, the C Scale was most strongly related, the R Scale next strongly related, and the H Scale generally unrelated.

Finally, the HCR-20 was related to an index of violence (past violent crimes). Items on the HCR-20 dealing with past violence were removed to avoid inflation of correlation co-efficients. Persons scoring above the median of the HCR-20 were significantly more likely than those scoring below the median to have previous violent convictions, previous assault charges, and juvenile records.


Summary

The conceptual risk assessment literature describes risk as multi-faceted (i.e., likelihood; severity; imminence; nature; targets) and calls for decision-makers to make decisions about these various facets. However, no research has been conducted to evaluate whether such highly specific judgments can be made reliably or accurately. This study evaluated whether highly specific judgments of violence could be made with reasonable reliability and incremental validity over more general judgments. For this study, the authors used the HCR-20 as the primary measure of violence risk. The sample consisted of 100 adults who had been found not criminally responsible by reason of insanity and were released from a maximum-security forensic institution into the community in 1996. The HCR-20 (Version) 2 was recoded from the original 175 participants described above because Version 1 had originally been used. The majority of the sample was male (n = 89). For this study, the definition of violence was categorized into three groups: any violence, physical violence, non-physical violence, and violence that resulted in criminal charges. Violence was measured from two sources: official criminal recidivism data and records of readmission to forensic psychiatric services.

The inter-rater reliability was calculated using ICCs. For the omnibus (general) structured clinical risk ratings on the HCR-20, the ICC1 was .61. Reliability was lower for more
specific judgments – from low to moderate for judgments of various severities of violence (ICC1 = .27-.37). Reliability was low to moderate for ratings of violence targets (ICC1 = .40-.47). Lastly, reliability was low to moderate for time frame of violence of up to one year (ICC1 = .31-.42).

Only a limited number of the specific judgments about future violence were capable of being evaluated due to low base rates, low reliability or lack of ability to collect outcome information. The short-term risk judgment produced a small but significant point biserial correlation with violence at 12 months ($r = .2; p = .02$). A partial point biserial correlation between short-term risk judgments and 12-month violence holding the general risk assessment constant was .06. Risk judgments of minor violence showed a significant point biserial correlation with non-physical violence ($r = .23; p = .012$). The partial point biserial correlation holding omnibus risk judgment constant was non-significant. Risk judgments of severe violence showed a significant point biserial correlation with physical violence ($r = .27; p = .003$). The partial point biserial correlation holding omnibus risk judgment constant was non-significant.

The findings showed that more specific judgments of various facets of violence risk that are called for in the literature were not made with as much reliability and accuracy as more general judgments.


**Summary**

The purpose of the study was to evaluate the relationship between confidence and accuracy of risk estimates made using a structured professional judgment (SPJ) and an actuarial approach. The impetus for the study was previous research (McNiel, Sandberg, & Binder, 1998) in which probabilistic clinical predictions of inpatient violence by civil psychiatric patients were influenced by clinicians’ confidence in their judgments.

The sample comprised 100 forensic psychiatric patients in western North America who had been found to be not criminally responsible for criminal offenses (previously reported on by Douglas, Ogloff, & Hart, 2003). The mean age at admission was 35.30 years (SD = 9.84) and participants primarily were single (67%), unemployed (93%), and had an admission diagnosis of schizophrenia (73.5%); 24.0% personality disorder; 18.4% mood disorder; 5.1% substance use/abuse disorder; 3% ‘other’). Most participants had a past violent charge (91.9%) and almost half had a past violent conviction (48.5%). The majority (79.0%) had a violent index offense.

The HCR-20 was coded archivally by two masters-level clinicians who were blind to outcome. Raters made actuarial predictions of risk (the sum of the HCR-20 items for each scale) and SPJ predictions of risk (final risk judgments of low, moderate, or high risk). They rated their confidence in their HCR-20 judgments on a 1-10 scale, where confidence was defined as “the rater has a feeling of certainty or reliance or trust about the correctness of the rating.” A median split was applied to create a low confidence group (who scored at or below the median) and a high confidence group (who scored above the median).

Four categories of violence were coded from criminal and hospital readmission records: (1) physical violence (physical contact by the perpetrator or use of a weapon); (2) nonphysical violence (verbal threats and fear-inducing behaviour); (3) criminal violence (violence that led to arrests or convictions); and (4) any violence (an omnibus category that included all violence).

A striking contrast emerged between effects of SPJs across the high and low confidence groups, with point biserial correlations ($r_{pb}$) and AUCs in the former typically being large and significant but in the latter being not significant. Correlations for any, physical, nonphysical, and criminal violence for the high confidence group were .62, .54, .48, and .43 and for the low confidence group were .14, .18, .10, and .03, respectively. AUC values for the any, physical, nonphysical, and criminal violence for the high confidence group were .86, .82, .82, and .84 and for the low confidence group were .58, .63, .58, and .52, respectively. Cox regression analyses, which control for time and uneven follow-up periods, yielded a nonsignificant model fit for the low confidence group using ‘any violence’ as the outcome criteria. However, in the high confidence group there was a roughly nine fold increase in the hazard of violence that occurred between low and moderate and between moderate and high risk ratings.

A similar set of analyses was carried out for the three actuarial judgments (one for each scale). In the low confidence group, all $r_{pb}$ and AUC values across the four violence categories were nonsignificant and generally small, whereas for the high confidence group the values generally were larger and were significant for the H- and C-scales (but not for the R-scale).

Indices of variability for scale scores and final risk ratings were highly comparable between the high and low confidence groups, which provided evidence against the possibility that the results could be attributed to differential
variance of the predictors between the two confidence groups. Several possible explanations for the strong relationship observed between confidence and accuracy are discussed.


Summary

This study tested the inter-rater reliability and criterion-related validity of structured violence risk judgments made with one application of the structured professional judgment (SPJ) model of violence risk assessment, the HCR-20 violence risk assessment scheme. Participants were taken from a larger, ongoing prospective study examining the predictive validity of the HCR-20. From the larger study, 116 of 175 patients released from forensic hospitalization between 1996 and 1997 were originally chosen to participate. The HCR-20 was completed on a random sample of 100 of the 116 forensic psychiatric patients. All of the 100 had been found not guilty by reason of insanity and were subsequently released into the community.

For this study, violence was operationally defined as actual, attempted or threatened physical harm to others. Acts of violence were divided into broad categories of: any violence, physical violence and non-physical violence. Raters were two masters-level clinicians. Raters gathered information from clinical-legal files of participants as they existed at time of discharge. Violence in the community was coded both from criminal records and clinical files after discharge from the hospital.

The mean HCR-20 total score was 24.7 (SD = 4.64). For the H-scale the mean was 14.4 (SD = 2.79), for the C-scale it was 4.68 (SD=2.02) and for the R-scale it was 5.88 (SD = 1.49). The ICC for the H-scale ranged from .41 (H4) to 1.0 (H7). For the total H scale it was .90. The ICC for the C-scale ranged from .34 (C5) to .69 (C3) (total C scale = .79), and for the R-scale, the ICC ranged from .01 (R5) to .54 (R3) (total R scale = .47). ICC for the HCR-20 total score was .85. As for agreement on final risk ratings, raters agreed on 70% of all cases, with no instances of low/high risk errors (ICC = .61).

AUC values for the HCR-20 structured clinical judgments (low, moderate, or high risk) were statistically significant for each outcome criterion. Effects for the HCR-20 clinical judgments were moderate to large in size, depending on the violence index (any violence, AUC = .69, p < .01; physical violence, AUC = .74, p < .01; non-physical violence, AUC = .68, p < .05). For the HCR-20 total score, the AUC for any violence was .67, p < .05; for physical violence was .70, p < .05 and for non-physical violence was .67, p < .05. For the H-scale, the AUCs were not significant. For the C-scale, the AUC for any violence was .68, p < .05; for physical violence it was .70, p < .05 and for non-physical violence it was .68, p < .05. For the R-scale, the AUC’s were not significant.

Kaplan-Meier survival analyses showed that persons judged to be at high risk were more likely to be violent, and to be so sooner than others. Cox regression analyses showed that HCR-20 risk ratings were most strongly related to violence, over and above actuarial scores.

The discussion section reiterates the findings and explores the implications of these results for using structured clinical judgments in risk assessments.


Summary

This presentation compared the H, C, and R scores of those patients who had been released by the Review Board to those who had not. While the H scale score did not differ between groups, C and R scale scores did. Among those discharged, the C scale score was 3.4, compared to 5.9 among those not released. Similarly, the R scale score was significantly lower among those released (4.0) compared to those not released (7.3).


Summary

This presentation reported the results of prospective analyses of the prediction of post-release violence among 103 released forensic patients followed for six months. The AUC value between “any aggression” and total score was .76. For H, C, and R, it was .60, .74, and .75. AUC values for PCL:SV total, Part 1, and Part 2 were .64, .57, and .66. For physical aggression, the AUC values were smaller: .57, .57, .60, and .61 for HCR-20 total, H, C, and R scale scores.
They were larger for PCL:SV total, Part 1, and Part 2 scores: .77, .75, and .70. As with Dernevik et al. (2002) and Müller-Isberner et al. (1999), Ross et al. (2001) suggested that risk management strategies could be responsible for the lower effects observed for more serious violence and HCR-20 scores. As with the other studies, however, this hypothesis remains untested. It is important to point out that the lower effects for more serious violence do not necessarily reflect a trend across studies, as other reports have failed to observe this (Douglas et al., 1999).


Summary

This is a chart review study of 80 men remanded to a secure forensic facility. The mean age at admission was 32.6 years (SD = 10.8). The majority of patients had been previously hospitalized in a psychiatric setting (77.5%), and most had previous charges or convictions for criminal offences (78.9%). Both the HCR-20 and the PCL-R averaged correlations just below $r = .30$ with several measures of later community violence. The HCR-20 was quite strongly related to subsequent re-admissions to the forensic hospital ($r = .38$) and to psychiatric hospitalizations ($r = .45$). The relationship of the PCL-R to these same outcomes was not as strong, at $r = .25$ and $r = .36$, respectively. However, whether these re-hospitalizations involved violence was not specified.

See also

Project and Scholarly Work


Summary

This study examined the utility of the HCR-20 and PCL: SV for predicting diversion noncompliance and reincarceration in a sample of defendants participating in diversion programs in New York. Both the HCR-20 and the PCL: SV have been validated in forensic, civil, and correctional samples, however this is the first study to examine the predictive utility of these instruments in the context of court-mandated treatment through diversion. The sample for this study consisted of 131 criminal defendants arrested for misdemeanor or felony charges who chose a mental health diversion option as opposed to incarceration. This study was part of a larger study for which 71 participants were interviewed and administered a battery of tests that included the HCR-20 and PCL: SV. To increase the sample size for this study, 60 additional files were randomly selected and the HCR-20 and PCL: SV were scored via file review only. The final sample included in this study was mostly male (68.7%). Participants had a mean age of 37.32 years (SD = 10.63). Most participants were African American (47.3%) or Hispanic (30.5%). In terms of clinical variables, 93.1% reported a history of mental health treatment and 92.4% reported a history of substance abuse. Of the sample, 29.8% of participants were diagnosed with major depressive disorder, 26.0% with bipolar disorder, 19.1% with schizophrenia, 8.4% with an anxiety-related disorder, 8.0% with delusional disorder, 6.1% with schizoaffective disorder, and 2.5% with other diagnoses. Prior to diversion placement, 72.5% of participants were charged with nonviolent offenses versus 27.5% of participants charged with violent offenses.

All interviews and file reviews were conducted by graduate students, who were trained in administration of the instruments. Recidivism data were coded from charts and incidents of diversion noncompliance were documented by participants’ case managers. Compliance with diversion was dichotomized into one category: diversion compliance versus diversion non-compliance. Diversion noncompliance was operationalized as the following: absconding from treatment program, failure to attend more than 2 days of outpatient program and/or weekly meetings with case manger, more than two incidents of drug or alcohol relapse and violent behavior (i.e. bodily acts of aggression against others, which is likely to cause harm to another person; clear and unambiguous threats of harm; and behavior likely to cause intimidation or fear). Outcome data were cumulatively recorded at 3, 6, and 12 months from diversion placement. Of the sample, 32.8% of participants were classified as noncompliant at 3 months, 42% at 6 months and 47.3% at 12 months. Regarding reincarceration, 27% were reincarcerated or out on a warrant at 3 months, 31.3% at 6 months and 38.9% at 12 months.

Inter-rater reliability of the HCR-20 and PCL: SV was examined on a subsample of approximately 40% of participants. ICC values for the H, C, R, Factor 1 and Factor 2 subscales were .93, .80, .61, .84 and .88 respectively. Item-level inter-reliability was also reported for each measure. The HCR-20 was found to be strongly positively correlated with the PCL: SV (r = 0.89, p < .01). For the sample as a whole, mean scores were: HCR-20 total M = 17.69 (SD = 6.01); H subscale M = 9.97, (SD = 3.59); C subscale M = 3.82, (SD = 2.05); R subscale M = 3.38, (SD = 1.83); PCL: SV M = 8.27 (SD = 4.59); Factor 1 M = 2.33, (SD = 2.36); and Factor 2 M = 2.33, (SD = 3.01). Mean scores on each of the measures were also reported for the following groups: diversion compliant, diversion noncompliant, non-reincarcerated and reincarcerated.

The authors conducted a logistic regression analysis to determine if the diversion noncompliance group could be classified by the HCR-20 and PCL: SV. Results indicated that both the HCR-20 and PCL: SV distinguished individuals who complied with diversion from those who did not comply; however, when controlling for the PCL: SV only the HCR-20 significantly predicted diversion compliance at 3-, 6- and 12-month follow-up periods. The same analyses were conducted for reincarceration as the criterion. Again, both the HCR-20 and PCL:SV distinguished individuals who were reincarcerated from those who were not. However, no individual scale significantly predicted the outcome when contributions of the individual scales were controlled for. Additional logistic regression analyses examining the subscales of the HCR-20 indicated that the R scale was the only significant predictor of diversion noncompliance at 6 and 12 months, whereas both the C and R scales were significant for noncompliance at 3 months. With regards to reincarceration, the H scale was the only significant predictor at 12 months.

Correctional Settings
(Includes Mentally Disordered Offenders)
The authors also conducted logistic regression analysis to explore whether the HCR-20 and PCL:SV predicted violence or recidivism at 12 months. No significant results were found. Low base rates for violence (9.9%) and recidivism (19.1%) likely contributed to these non-significant findings.

The authors report the findings of receiver operating characteristic analyses. With regards to noncompliance, AUC values were .79, .73, .71, .74, .72, .60, and .74 for the HCR-20 total, H, C, R, PCL: SV total, Factor 1 and Factor 2 scores, respectively. With regards to reincarceration AUC values were .71, .68, .63, .66, .68, .59, and .69 for the HCR-20 total, H, C, R, PCL: SV total, Factor 1 and Factor 2 scores, respectively. Regardless of outcome, PCL: SV Factor 1 scores were not significantly predictive.

This study also examined whether the HCR-20 added incremental validity to the PCL:SV, even if H7 (the psychopathy item) was removed. To determine if the results would remain the same, a sequential logistic regression analysis with diversion noncompliance as the criterion variable was conducted with the PCL:SV total score entered first, followed by the HCR-20-7. Comparison of log likelihood ratios of models with and without the addition of HCR-20-7 showed significant improvement when HCR-20-7 was added to the equation, with HCR-20-7 being the only variable that contributed significantly to the prediction. When the order was reversed, the PCL:SV failed to contribute significantly to the predictive power of the HCR-20 alone.

Results provided support for the use of the HCR-20 and PCL:SV in the context of diversion, although the HCR-20 proved to be superior to the PCL:SV in predicting diversion noncompliance and reincarceration.

SEE ALSO


Project and Scholarly Work


Summary

This study examined the sociobiographic, criminological, and clinical characteristics of 58 men for whom preventative detention had been ordered in two German federal states. For all participants, the PCL-R and HCR-20 were completed using a semi-structured clinical interview and/or collateral file review. If an offender had committed a sexual offense the SVR-20 and Static-99 were also coded.

The majority of the sample (91.4%) was German. The mean age of the individuals at the time of assessment was 50 years ($SD = 11.41$). Alcohol abuse/disorder (41.4%), pedophilia (20.75%) and antisocial personality disorder (81.0%) were the most common diagnoses in the sample. To determine whether there were any significant differences between participants based on index offense type the sample was divided into the following groups: sexual offenders ($n = 35$), violent offenders ($n = 14$) and Robbêrs ($n = 9$).

Means for the sample as a whole and each of the three subgroups were reported. Mean total score on the PCL-R was 23.9 ($SD = 6.18$) for the entire sample. Approximately 10.3% of the sample scored 30 or higher on the PCL-R. The group of violent offenders had higher values for Factor 2 ($M = 14.3$, $SD = 3.20$) compared with both other groups (sex offenders $M = 10.9$, $SD = 4.01$ and Robbêrs $M = 13.8$, $SD = 3.87$). Violent offenders also had the highest total PCL-R total scores ($M = 26.6$, $SD = 5.49$ vs. $M = 22.26$, $SD = 6.28$ vs. $M = 26.33$, $SD = 5.00$). Values for Factor 1, however, were higher for Robbêrs compared to the other two groups ($M = 10.67$, $SD = 2.83$ vs. $M = 8.97$, $SD = 3.12$ vs. $M = 10.64$, $SD = 2.73$). Mean total score on the HCR-20 was 25.84 ($SD = 6.23$). Similar to the PCL-R, violent offenders had higher total scores ($M = 29.21$, $SD = 5.10$) compared with both other groups (sex offenders $M = 24.40$, $SD = 7.22$, Robbêrs $M = 26.22$, $SD = 4.47$). For sex offenders mean total scores were 22.06 ($SD = 5.09$) and 6.26 ($SD = 1.85$) on the SVR-20 and Static-99, respectively. The authors’ note that because of the small size results may not be representative of the larger population of individuals under preventative detention.

Project and Scholarly Work


Summary

The HCR-20 violence risk assessment scheme (Version 2) and the Psychopathy Checklist: Screening Version

---

154
(PCL:SV) were coded on a sample of 41 male inmates from two Swedish maximum security prisons. The two coders were Ph.D. and M.D. level clinician-researchers. This was a prospective study of violence within the correctional institution over an eight month period. The HCR-20 and PCL:SV were coded by use of both file review and clinical interview. The R Scale of the HCR-20 was coded using the “In” option as explained in the manual (Version 2).

The mean age of the participants was 35, and the mean length of incarceration at time of assessment was three years. All participants had a personality disorder (mostly antisocial). Of the 41, 27 were incarcerated for homicide, and 14 for other violent offences. The sample was highly psychopathic, with 30 of 41 inmates being classified as psychopaths.

Eight of the 41 (19.5%) inmates were violent in the prison. The C Scale, R Scale, HCR-20 Total Score, PCL:SV Part 2, and PCL:SV Total Score differentiated between the violent and non-violent groups. The HCR-20 Total score was 33.4 in the violent group, and 24.6 in the non-violent group. All HCR-20 R Scale items were significantly greater among the violent group than the non-violent group. The H Scale was not predictive of violence, except for Item H10. In the group of 30 psychopaths, the R Scale and HCR-20 Total score were significantly higher in the violent inmates. Four of the five R Scale items were higher in the violent psychopaths compared to the non-violent psychopaths.

Results imply that the HCR-20 (and the PCL:SV) are predictive of violence by inmates within correctional institutions. Even among a sub-group of psychopaths, the HCR-20 distinguished between violent and non-violent inmates. The authors comment that the H Scale was not predictive in this sample because inmates (being maximum security violent inmates) were homogeneous with respect to most historical factors. The Clinical and Risk Management factors did, however, provide for a means of separating violent from non-violent inmates. These results are consistent with those of Strand et al. reported above. The results of the study, though limited by a small sample, provide support for the importance of risk management concerns for high-risk violent offenders.

This research focused on whether the use of comprehensive risk assessments and subsequent risk management could prevent institutional violence in a maximum security correctional institution in Sweden. More specifically, the authors investigated whether increasing staff members’ knowledge of risk factors and appropriate risk management strategies would affect the rate of violence on a ward.

All participants (N = 47) were offenders with a violent criminal history and who were incarcerated at some time between October 1999 and June 2002 on one ward of the institution. The authors noted that this particular ward (A-ward) is not meant to house any specific type or class of offender and that it is not known to have an elevated level of psychiatric problems among prisoners relative to other wards. However, many of individuals housed on A-ward were transferred there for misbehaviour. Participants on average were 32 years old and most frequently were diagnosed with antisocial personality disorder (n = 29). No participants were diagnosed with a thought disorder. Eighteen participants had PCL:SV scores of 18 or above, and psychopathic offenders had significantly higher HCR-20 scores compared with nonpsychopathic offenders (p = .000, Mann-Whitney U-test).

Prior to the study’s commencement, staff received training on risk assessment, common risk factors for violence, foundational knowledge about the HCR-20 and PCL:SV, and adequate risk management strategies for different types of mentally disordered offenders. Ongoing training also was provided throughout the study. Two clinicians collaboratively completed a risk assessment on each participant that included the Swedish versions of the HCR-20 and the PCL:SV. Following completion of the risk assessment, results were discussed with staff members and a risk management strategy for the individual was developed.

The HCR-20 was readministered to roughly one third (n = 13) of the sample. The average length of time between the two HCR-20 assessments was 12 months (range: 3-24 months). Comparison of HCR-20 scores between the follow-up group of 13 and the entire study group of 47 revealed no significant changes. However, there was a significant decrease in incidence of violence (from an average of 14 violent incidents per year between 1993 and 1998 to an average of 5 violent incidents per year during the study period of 1999-2001; p = .024, Mann-Whitney U-test). This 64% reduction is in stark contrast to all other wards in the prison, in which there were no decreases in incidence of violence during the study period.

The discussion section advances possible explanations for the results. The authors highlight the importance of incorporating into a risk assessment protective factors.

Project and Scholarly Work


Summary
which they noted can reduce violence even when important risk factors do not decrease.

Project Description


Summary

The Prisoner Cohort study originally began in order to investigate the predictive accuracy of several risk assessment instruments with a main goal of identifying the best instruments for use in the UK. The project began as a part of the Dangerous and Severe Personality Disorder (DSPD) program. Offenders were included in the study if they were serving a sentence of two or more years for a violent or sexual offence, were at least 18 years old, and had less than one year of their sentence left to serve. A stratified random sample was recruited resulting in a sample of 1396 males. Participants were mostly White (79%), while a minority was Black (15%). Their mean age was 30.8 years (SD = 11.4). The breakdown of index offences were as follows: 45% Robbery, 24% major violent offence, 23% major sexual offence, and 18% minor violence.

Risk assessments were completed by trained interviewers during the 6 to 12 month period before each offender’s expected release date. Offenders were assessed using the HCR-20, the PCL-R, the VRAG, the Risk Matrix 2000 Violence Scale (RM2000V), and the Offender Group Reconviction Scale (OGRS-2). Outcome data was based upon official reconviction data obtained from the Police National Computer system. The average length of the follow-up period was 1.97 years, with a range from six days to 2.91 years. Outcome data were available for 1353 participants.

A total of 41% of the sample were reconvicted in the follow-up period for various charges including breach (19%), theft (17%), driving (12%), minor violence (8%), drug related offences (8%), or other crimes. The mean scores on the HCR-20 were 19.1 (SD = 7.80), 11.1 (SD = 4.56), 3.39 (SD = 2.15), and 4.50 (SD = 2.55) for the total score, H, C, and R scales, respectively.

All of the instruments used were found to be predictive of violent reconvictions: HCR-20 total score (AUC = .638), H scale (AUC = .676), C scale (AUC = .643), R scale (AUC = .592), OGRS-2 (AUC = .724), VRAG (AUC = .700), PCL-R (AUC = .639), and RM2000V (AUC = .687). The authors also report AUCs for reconvictions of theft, drug offences, and any offence. For these three outcomes all of the instruments were also predictive. Also reported is reconvictions of Robbery for which only the OGRS-2, VRAG, HCR-20 H scale, and RM2000V were predictive. The instruments were compared directly by fitting them in a regression model. Entering all of the various instrument total scores and the HCR-20 scale scores in the regression model, the OGRS-2 was found to outperform the other instruments for all of the outcome offences. Nonetheless, all of the instruments were significantly predictive in the models with lower betas. The OGRS yielded betas ranging from .65 to .94. The VRAG (β = .48 -.81), RM2000V (β = .51 -.79), and H scale (β = .43 -.73) tended to perform slightly below the OGRS-2. The PCL-R (β = .26 -.56), C scale (β = .08 -.48) and R scale (β = .13 -.40) tended to perform at the lowest level according to the regression model fitting.

Predictive validity at the item level was also analyzed for each of the instruments independently by entering all the items in a logistic regression model. Only the results for the HCR-20 are reported here. All of the H scale items related to at least one of the outcome offences. C2 (negative attitudes) and C4 (impulsivity) were the only C scale item found to be predictive at the item level. R1 (plans lack feasibility), R2 (exposure to destabilizers), and R4 (non-compliance with remediation attempts) were also found to be predictive.

Various additional analyses for the other risk assessment measures are not reported here to save space. The authors also discussed various results based on separating the sample into those that met the definition of DSPD and those who did not.

Scholarly Works


Summary

The current study examined gender differences in risk assessment instruments based on data from the Prisoner Cohort study (discussed above) and a similar sample from the same settings consisting only of women (N = 391). All procedures were identical for the female samples as those discussed above. The follow-up period for the female sample was slightly shorter with an average of 1.40 years (SD = 0.86).
For the female sample, mean scores for the HCR-20 were as follows for the total score, H, C, and R scales, respectively: 20.2 (SD = 7.48), 12.1 (SD = 4.29), 3.11 (SD = 2.11), and 5.03 (SD = 2.42). Mean difference tests revealed significant differences between males’ and females’ HCR-20 total scores and scale scores with females having higher scores on all scales expect the C scale. Mean differences were also found for the OGRS and PCL-R with males having higher scores.

With respect to predicting violent reconvictions, the effect sizes were slightly larger for females compared to males. For females, the AUCs were .70, .73, .69, and .59 for the total score, H, C and R scales, respectively, compared to .67, .66, .64, and .59 for each of the respective scores for the male sample. With respect to predicting acquisitive reconvictions, the effect sizes were slightly larger for the males compared to the females. For males, the AUCs were .69, .70, .63, and .61, for the total score, H, C and R scales, respectively, compared to .61, .62, .60, and .55 for each of the respective scores for the female sample.

Multivariate regression analyses were conducted to determine the independent effects of various instruments and scales. Using violent reconvictions as the dependent variable, each of the HCR-20 scales was entered in the regression model. For men, the results revealed that the H scale outperformed the C scale which in turn outperformed the R scale. Using acquisitive reconvictions as the dependent variable, the H scale outperformed the C scale which in turn outperformed the R scale. The trend for women was less clear; however, the H scale generally performed best.

This study reports various results for the other risk assessment instruments. Additional results are also reported for any reconviction. The authors discussed various implications of the results and limitations of the present design.


**Summary**

Reporting on the same data from the Prisoner Cohort study, the present paper focused on examining the independent predictive ability of individual items within the risk assessment instruments. The authors first report point biserial correlations between the HCR-20 and violent reconvictions \( (r = .31, .31, .26, \text{ and } .18 \) for the total score, H, C, and R scales, respectively). Item level correlations revealed that all but three items (H6, C3, and R3) in the HCR-20 were predictive of violent reconviction.

Next, stepwise regression was used to determine which items were independently predictive of violent reconvictions. This analysis revealed that 8 of the 20 items in the HCR-20 were independently predictive: H2 (young age at first violence), H5 (substance use problems), H8 (early maladjustment), H10 (prior supervision failure), C2 (negative attitudes), C4 (impulsivity), R2 (exposure to destabilizers), and R4 (noncompliance with remediation attempts).

The authors then examined whether a new aggregate score that consisted of only the independently predictive items yielded larger effect sizes than the full total score. For the HCR-20, the AUC for the 8 items scale was .69 compared to .67 for the full total score. This difference was not statistically significant.

The authors report similar analyses for the PCL-R and VRAG. Totalling only the independently predictive items of the instruments did not result in significantly better predictive accuracy. The authors discussed the possible limits to predictive accuracy including a “glass ceiling effect” and implications of the results in terms of inclusion of items in risk assessment instruments.


**Summary**

Previous studies that have compared logistic regression (LR), classification and regression tree (CART), and neural networks (NNs) models for their predictive validity have shown inconsistent results in demonstrating superiority of any one model. The current study tested the three models in a prospective sample of 1125 UK male prisoners followed up for a mean of 3.31 years after release. Items of the HCR-20 were used as predictors and violent reconvictions as the criterion variable. Participants in the study had a mean age of 30.87 years (SD = 11.3 years). A majority (78.9%) was white British and had a diagnosis of personality disorder (74.0%). Twenty-eight percent of the sample was reconvicted for violent offenses during the follow-up period.

Violent reoffenders scored significantly higher on the HCR-20 total than the non-violent reoffenders. The historical subscale demonstrated the highest differential effect amongst the three subscales. For the prisoners that violently reoffended, mean scores on the HCR-20 total, historical, clinical and risk management scales were 22.21 (SD = 6.55), 13.07 (SD = 3.38), 4.08 (SD = 2.04), and 5.06 (SD = 2.58). For the prisoners that did not violently reoffend, mean...
scores on the HCR-20 total, historical, clinical and risk management scales were 17.88 (SD = 7.84), 10.37 (SD = 4.69), 3.15 (SD = 2.11) and 4.36 (SD = 2.48).

Correlation coefficients between each of the individual items of the HCR-20 and violent recidivism were presented. Six items were not significantly correlated with the outcome: relationship instability (H3), major mental illness (H6), active symptoms of major mental illness (C3), unresponsive to treatment (C4), lack of personal support (R3), and noncompliance with remediation attempts (R4). Young age at first violent incident (H2) was most strongly correlated with the outcome (r = 0.24), followed by prior supervision failure (H10; r = 0.21).

To compare LR, CART, and NNs models for their predictive validity, traditional ROC analysis was conducted based on the total scores for the HCR-20. A multi-validation procedure was used to reduce sampling error in reporting the predictive accuracy. The total sample was divided randomly into four subsets and the authors combined these subsets into four different validation samples. The four subsets were: Subset A (N = 306, base rate = 27%), Subset B (N = 306, base rate = 28%), Subset C (N = 307, base rate = 28%) and Subset D (N = 306, base rate = 29%). The efficacy of the HCR-20 for violent reconviction in the four different validation samples ranged from .58 to .70 in terms of AUC values and .57 to .67 in terms of overall accuracy. Accuracy of the three models varied between .59 and .67, with an overall AUC range of .65 to .72. Although the performance of NNs was slightly better than that of LR of CART models, it did not demonstrate significant improvement.

The authors concluded that in general, the three models demonstrated similar levels of accuracy in predicting violent reconviction. They note that the efficacy of the HCR-20 was lower than that reported in some studies and this may be because of the heterogeneity of participants.

See Also

Coid, J. (2009, June). All predictive instruments have the same predictive abilities. In J. Coid (Chair), Deconstructing risk assessment for violence. Symposium conducted at the annual conference of the International Association of Forensic Mental Health Services, Edinburgh, Scotland.


Project and Scholarly Work


Summary

This was a retrospective follow-up of 250 male adult prisoners released from a Scottish prison who had been randomly selected from the larger population. Mean age was 26.8; almost all were Caucasian; roughly half were unmarried (46%); offences included drugs (36%), assault (24%), theft and break and enter (20%), homicide (8%), weapons (7%), kidnapping (3%), and sexual offences (2%). Prisoners had received earlier comprehensive correctional assessments, using interview and file procedures. Psychopathy was one of the constructs that was assessed. The HCR-20 later was coded from files. The researchers were unable to code C1, C5, R2, R4, and R5. As they commented, this may have underestimated the predictive accuracy of the HCR-20. Violence was measured with the MacArthur Community Violence Instrument. The authors carried out comprehensive analyses of the HCR-20, VRAG, and PCL-R, down to the item (this is a 128 page report!).

Interrater reliability, using ICC, on a subsample of 60 prisoners, was described as “acceptable and … consistent with values obtained in other studies” (p. 30). Values were as follows: HCR-20 Total Scale Score (.92); Historical Scale Score (.92); Clinical Scale Score (.74); Risk Management Scale Score (.70). Lower indices for C and R may stem from coding files only.

Survival analyses of individual HCR-20 items led the researchers to conclude that “Overall, these analyses indicate that the vast majority of these items have some potential utility as predictors” (p. 46). Survival analyses indicated that the Total H Scale Scores were related to each outcome measure (including both violent and non-violent recidivisms); the C Scale was related to the two violence outcome measures, and the R Scale was related to the general recidivism outcome measures, but only weakly to violent outcomes. Recall that only two of five R Scale items were coded.

Cox proportional hazard model analyses revealed the $e^b$ for re-incarceration for violence was 1.20 (exponentiated value of the model parameter; an effect size indicating the change in hazard rate as a function of a unit change in the total HCR-20 score). This means that for every 1-point increase in the HCR-20, the hazard for violence increases 20%. For a 5-point increase, the hazard increases by 200% (100% +...
5x20%). Using the H Scale alone, these analyses revealed that 14% of those with a mean score (10.9) were returned to prison for violence, 4% with a score of 5, and 72% with a score of 20. This was greater discrimination than achieved for the PCL-R or VRAG. When compared directly, the H Scale was the strongest predictor of return to prison for violence; the VRAG was stronger for general recidivism and violent recidivism that did not lead to re-imprisonment (presumably less serious violence). Cooke et al., based on these and other analyses, interpreted the findings as suggestive of greater specificity for the prediction of violence, in general, and likely serious violence, compared to community recidivism, for the H Scale. HCR-20 Total Scale Score (.25); H Scale Score (.24); C Scale Score (.23); R Scale Score (.10).

This study used the German version of the HCR-20. It drew from data from the Berlin CRIME study, which was a longitudinal study of 397 criminal offenders released from prison in 1976. The HCR-20 was coded on 200 randomly selected prisoners from these 1976 files. The researchers were unable to code H7 – the Psychopathy item. Because the sample was random, index offences mainly were property offences, though there were some serious violent offences as well. The inmates averaged 31.42 (SD = 5.40) years in age, and had 6.46 (SD = 4.72) previous convictions. More than half (65%) had committed previous violent offences. Many had alcohol (56%) problems; fewer (17%) had drug problems.

Based on a subsample of 30 offenders, intrarater reliability (Kendall’s Tau) was .80 for the HCR-20 Total Scale. Spearman’s Rho was .731 for the C Scale and .930 for the H Scale. Interrater reliability was not reported for the R Scale. Over the 20 year follow-up, the correlation between the HCR-20 and violent recidivism was as follows: HCR-20 Total Scale Score (.25); H Scale Score (.24); C Scale Score (.23); R Scale Score (.10).


### Summary

The ability of the HCR-20, PCL-R, and LSI-R to predict officially recorded convictions was examined among male prisoners in Berlin (N = 307) who survived at least ten years at risk following release from prison. All instruments were coded retrospectively by psychologists who were trained how to use them. The mean length of follow-up after release was 19.6 years (SD = 1.74 years). At the time the study started in 1976, participants’ mean age was 30.52 years (SD = 5.38 years; range = 21-42 years).

Descriptive information for the measures was as follows: HCR-20 Total (M = 16.52, SD = 6.31, α = .84); HCR-20 Historical scale (M = 8.32, SD = 3.51, α = .69); HCR-20 Clinical scale (Md = 3, range = 8, α = .49); HCR-20 Risk Management scale (Md = 5, range = 10, α = 80); PCL-R Total (M = 12.03, SD = 4.70, α = .71); PCL-R Factor 1 (Md = 3, range = 12, α = .59); PCL-R Factor 2 (Md = 6, range = 17, α = .71); LSI-R Total (M = 24.65, SD = 7.35, α = .84). The intercorrelations between the three measures were high: HCR-20/LSI-R (.80); HCR-20/PCL-R (.76); LSI-R/PCL-R (.61).
Interrater reliability was examined on a sub-sample of participants ($n = 30$) using two independent raters. Values were as follows: HCR-20 Total (ICC = .91; $r = .91$; 95% CI = .83-.96); HCR-20 Historical scale (ICC = .92; $r = .92$; 95% CI = .83-.96); HCR-20 Clinical scale (ICC = .82; $r = .83$; 95% CI = .65-.91); HCR-20 Risk Management scale (ICC = .78; $r = .79$; 95% CI = .59-.89); PCL-R Total (ICC = .94; $r = .94$; 95% CI = .88-.97); PCL-R Factor 1 (ICC = .80; $r = .81$; 95% CI = .63-.90); PCL-R Factor 2 (ICC = .89; $r = .89$; 95% CI = .77-.94); LSI-R Total (ICC = .93; $r = .93$; 95% CI = .86-.97).

For each measure, predictive validity data were collected for different lengths of time at risk. All of the values that follow correspond to a criterion of violent reconviction. Predictive validity for the HCR-20 Total score was: 2 years at risk ($r = .21$, $p < .001$); AUC = .75, $p = .06$, 95% CI = .61-.85); 5 years at risk ($r = .28$, $p < .001$; AUC = .71, $p = .04$, 95% CI = .63-.80); total time at risk ($r = .28$, $p < .001$; AUC = .69, $p = .04$, 95% CI = .62-.76). For the HCR-20 Historical scale: 2 years at risk ($r = .18$, $p < .05$); 5 years at risk ($r = .26$, $p < .001$); total time at risk ($r = .27$, $p < .001$). For the HCR-20 Clinical scale: 2 years at risk ($r = .18$, $p < .05$); 5 years at risk ($r = .22$, $p < .001$); total time at risk ($r = .25$, $p < .001$). For the HCR-20 Risk Management scale: 2 years at risk ($r = .16$, $p < .01$); 5 years at risk ($r = .18$, $p < .01$); total time at risk ($r = .17$, $p < .01$).

Predictive validity for the PCL-R Total score was: 2 years at risk ($r = .14$, $p < .05$; AUC = .67, $p = .06$, 95% CI = .55-.78); 5 years at risk ($r = .25$, $p < .001$; AUC = .70, $p = .04$, 95% CI = .63-.78); total time at risk ($r = .27$, $p < .001$; AUC = .68, $p = .04$, 95% CI = .61-.75). For Factor 1, values were: 2 years at risk ($r = .01$, $p = ns$); 5 years at risk ($r = .13$, $p < .05$); total time at risk ($r = .16$, $p < .05$). For Factor 2, values were: 2 years at risk ($r = .15$, $p < .01$); 5 years at risk ($r = .24$, $p < .001$); total time at risk ($r = .23$, $p < .001$).

Predictive validity for the LSI-R Total score was: 2 years at risk ($r = .15$, $p < .01$; AUC = .68, $p = .06$, 95% CI = .56-.79); 5 years at risk ($r = .21$, $p < .001$; AUC = .67, $p = .04$, 95% CI = .60-.75); total time at risk ($r = .20$, $p < .001$; AUC = .64, $p = .04$, 95% CI = .57-.71).


Summary

An adapted German version of the HCR-20, the LSI-R, and the PCL-R were scored from data collected as part of a longitudinal study on the biographical development of criminal offenders conducted in former West Berlin. The original sample comprised every fourth adult male who began serving a sentence between February and May 1976 ($N = 397$). No selection was made regarding the type of crime (with the exception of pure traffic offense), the length of sentence, or the level of security. Analyses were based on the subset of offenders who survived for ten years or longer after release ($n = 307$). Participants’ mean age at the time of admission was 29.83 years ($SD = 5.35$; range: 21-45).

Psychologists blind to recidivism completed the risk assessments retrospectively based on the time when participants were released from prison. Assessments were based on data gathered during the course of a basic examination, information from the prisoners’ personal files accumulated throughout the prison sentence, and the offender’s social situation at the time of release. Ratings on HCR-20 Historical scale items primarily were based on information gathered during the basic examinations at the beginning of the sentence, whereas items on the Clinical and Risk Management scales primarily were rated based on behavioural descriptions available in the prison personal files. Recidivism data were gathered from criminal records available at periods of two, five, and ten years post-release.

Agreement among raters (ICC) regarding the instruments’ total scores was high: LSI-R (r = .93); HCR-20 (.91); PCL-R (.94). Agreement also was high for the HCR-20 scales: Historical (.92); Clinical (.82); Risk Management (.78). Internal consistency of the instruments’ total scores were: LSI-R (α = .84); HCR-20 (α = .84); PCL-R (α = .71). Mean values of the instruments’ total scores were: LSI-R (24.65; $SD = 7.35$); HCR-20 (16.52; $SD = 6.31$); PCL-R (12.03; $SD = 4.70$). Intercorrelations between the instruments generally were high: LSI-R and HCR-20 ($r = .80$); LSI-R and PCL-R ($r = .61$); HCR-20 and PCL-R ($r = .76$).

With the exception of the PCL-R, the accuracy of predictions for general recidivism decreased gradually as the length of follow-up increased. In contrast, the accuracy for predictions of violent recidivism (the specific operationalization of which was not defined) improved as the length of observation increased. The correlations between instruments’ total scores and violent crime for periods of two, five, and ten years at risk, respectively, were: HCR-20 (.21; .28; .31); LSI-R (.15; .21; .23); PCL-R (.14; .25; .32). The correlations between instruments’ total scores and re-imprisonment for periods of two, five, and ten years at risk, respectively, were: HCR-20 (.37; .34; .31); LSI-R (.41; .34; .29); PCL-R (.31; .32; .34). Differences in predictive accuracy between the instruments were not statistically significant.

Additional analyses indicated that the instruments’ predictive accuracy was dependent on offenders’ demographic, criminological and psychopathological characteristics. The author concluded that each risk assessment measure was applicable to German criminals.
and required only a few adaptations to be used in German, and that the instruments demonstrated levels of predictive accuracy that were comparable to the values reported in the literature (for non-German samples).

---

**Project and Scholarly Works**


**Summary**

This study compared the predictive validity of five indices of violence risk – the HCR-20, the VRAG, the VORAS, the PCL:SV, and the PCL-R. Participants were 188 male offenders released from federal corrections institutions to supervision in Western Canada. Participants were selected based on known outcome status after release. This status was as follows: violent recidivism (N = 93) or no violent recidivism after release (N = 95). The follow up period for this study ranged from 6 to 11 years.

Inter-rater reliability was good to excellent for all measures except for the HCR-20 structured final judgment ratings which can be considered fair to moderate. Even with the lower inter-rater reliability, there were no low/high disagreements with this item.

Point-biserial correlations between risk assessment measures and violent recidivism showed that with the exception of the H-scale, the HCR-20 total and sub-scales produced correlations of approximately .50. The aforementioned H-scale showed a correlation of .36. The AUCs for the HCR-20 were approximately .80 (up to .82) except for the H-scale (.72). Partial point-biserial correlations were also conducted with the HCR-20 total score, with the correlation to violent recidivism dropping from .51 to .25 after controlling for the VRAG, PCL-R and the VORAS. Other measures also showing a significant positive point-biserial correlation after controlling for other measures were: HCR-20 C and R scales and structured final judgment, the VRAG, VORAS B, and Cooke and Michie’s third factor. The other measures were either not significantly related or were negatively related to the outcome.

Binary logistic regressions were conducted to directly compare the measures. Using forward conditional entry procedures only the HCR-20 total score entered the model. Using direct entry procedures the HCR-20 total score, VRAG total score, and the VORAS total score entered the model. Using subscales instead of total scores, direct entry showed that the HCR-20 C-scale, VORAS A (negatively) and VORAS B were significant predictors. Using forward entry, the HCR-20 C-scale, VORAS A (negatively) and VORAS B were again significant predictors. Using the HCR-20’s and other measures’ final risk judgments showed that the HCR-20 structured clinical final risk judgment, the VRAG’s actuarial categorical system, and the VORAS actuarial final risk score were significantly predictive.

The discussion section explores the implications of the results from this study in regards to the use of the five measures analyzed here.


**Summary**

The HCR-20 violence risk assessment scheme was coded in a sample of 72 Canadian, male, federally-sentenced, maximum security offenders who had been referred to a regional health centre of the Correctional Services of Canada. The concurrent validity of the HCR-20 was assessed through comparison to other instruments (the Psychopathy Checklist - Revised; Violence Risk Appraisal Guide) and to the presence of several past indexes of violent and antisocial behaviour. Only the H and C scales could be coded because no offenders had yet been released. This was a postdictive study.

The interrater reliability of the H and C combined scores was .80. Correlations between the number of previous violent charges and the H scale, C scale, and their combination ranged from moderate to large. The Historical scale correlated at r = .50 with previous violence (with the “previous violence” item removed from the H scale), the Clinical scale at r = .30, and the combined total at r = .44. The VRAg correlated at r = .20 with previous violence, and the PCL-R’s correlation with past violence was r = .41.

Scores above the median of the HCR-20 increased the odds of the presence of various measures of past violence and antisocial behaviour by an average of four times. The main limitations of this research were a small sample and a retrospective design.

**See also**

Project and Scholarly Works


Summary

This study compared the validity of the Historical component of the HCR-20, the Psychopathy Checklist: Screening Version (PCL:SV; Hart, Cox, & Hare, 1995) and the Violence Risk Appraisal Guide (VRAG; Harris, Rice, & Quinsey, 1993). This research used 87 adult mentally disordered patients in a medium secure unit. An independent researcher, who was blind to the score on the risk assessments, collected data on violent outcome measures during the first 12 weeks after admission for all 87 subjects.

For this study, violence was operationally defined as actual, attempted or threatened harm to others. Violence episodes were dichotomized into two levels. Level 1 involved physical assault against a person or any violence resulting in injury to a person. Level 2 included other aggressive acts such as threats or property damage.

On the risk assessment tools, those with any violent outcome had significantly higher scores on all measures, with the PCL:SV showing the most significant group difference ($t = 4.77; p < .0001$). The VRAG ($t = 3.6; p < .001$) and the H-10 ($t = 3.2; p < .001$) also showed a significant group difference.

Using ROC analyses for the prediction of any type of violence, the AU C's produced for the PCL:SV total score ranged from $.76 (p < .001)$ for any and physical violence to $.74 (p < .01)$ for level 1 violence. The AU Cs produced for the VRAG total score ranged from $.71 (p < .01)$ for any and physical violence to $.64 (p < .01)$ for level 1 violence and the AU Cs produced for the H-10 total score ranged from $.70 (p < .01)$ for any and physical violence to $.66 (p < .01)$ for level 1 violence.

Using a cut-off score of 18 (recommended manual cut-off) on the PCL:SV, the odds ratio for any violence was 4. Using a cut-off score of 5 (sample mean) on the VRAG, the odds ratio was 3.75 for any violence. Using a cut-off score of 12 (sample mean) on the H-10, the odds ratio was 3 for any violence.

The PCL:SV total score correlated significantly with the VRAG score ($r = .81, p < .001$) and with the H-10 score ($r = .8, p < .01$). The VRAG score and the H-10 score also correlated significantly with each other ($r = .83, p < .001$).

Forward stepwise logistic regression showed that only the PCL:SV total score contributed significantly to the prediction of any violence ($\chi^2 = 20.05, p < .001$).

Implications for research on risk assessment, as well as the clinical assessment and management of violence, are discussed.


Summary

This research studied the differences on the HCR-20, VRAG and PCL-R between groups of men who committed single acts of violence vs. men who have a history of committing multiple acts of violence. The authors point out that there may be difficulty in predicting future actions of men who have only committed one single act of violence due to the fact that that risk assessment research is predicated on the assumption that static indicators of past behaviour can be used to predict future behaviour. Subjects were male mentally disordered offenders, 21 years of age or older with one or more convictions for violence or sexual violence.

In terms of the HCR-20, there were differences between groups on many individual items from the three subscales. On the H-scale, there were significant differences on H2, H7, H8, H9, and H10, with repeat offenders scoring higher on all of these items. On the C-scale, there were significant differences on C2, C3 and C5, with repeat offenders scoring higher on C5 and C2 and lower on C3. On the R-scale, there were significant differences on R1, R2, R3 and R4, with repeat offenders scoring higher on all of these items. In terms of the PCL-R, repeat offenders scored significantly higher on the total score and on the factor two scale. In terms of the VRAG, repeat offenders scored significantly higher on the total score.

Summary

This research project investigated the comparability of outcomes when the HCR-20 was given in single administrations or in group administrations. This study sampled from a group of 85 male mentally disordered offenders in a personality disorder unit of a major forensic psychiatric facility in England. Descriptive statistics for group administrations were as follows: HCR-20 Total score, N = 68, M = 27.88 (SD = 5.01); H-score, N = 73, M = 15.15 (SD = 2.53); C-scale, N = 84, M = 5.27 (SD = 2.2) and R-scale, N = 80, M = 7.5 (SD = 2.27). Descriptive statistics for single administrations were as follows: HCR-20 Total score, N = 36, M = 27.33 (SD = 4.76); H-score, N = 36, M = 14.75 (SD = 2.45); C-scale, N = 36, M = 5.53 (SD = 1.95) and R-scale, N = 36, M = 7.06 (SD = 1.67). None of these scores for the group or single administrations were significantly different from each other.

Correlations (rho) between group and single administrations were mostly significant. HCR-20 total (r = .428, p < .05), H-scale (r = .654, p < .01) C-scale (r = .503, p < .05) and R-scale (ns).

The authors concluded that if there are no differences between single and group administrations of the HCR-20, then why not make use of group administrations.

Project and Scholarly Works


Summary

This study was not a validation study of the HCR-20, but rather an analysis of the characteristics of hate-motivated violent offenders in California, and comparison of this group to other offenders. The HCR-20 was used as a measure of violence risk for this purpose. The HCR-20 was coded from files of 58 hate-motivated criminals. Most of the offenders were male (n = 53, 91.4%), with a mean age of 24.5 (SD = 8.07). Most offenders were Latino (48.3%), followed by Euro-Caucasian (32.8%) and African American (15.5%). Close to half were unemployed (45.5%). The majority of the offenders had substance use problems (58.6%), and many (22.4%) had received past psychiatric treatment. The vast majority of offenders had previous convictions (87%) and a history of violence (60%). All offenders had violent, hate-related index offences.

Analyses on the HCR-20 included correlations with the PCL-R, with violent offences, and with the Cormier-Lang crime severity scales. The HCR-20 did not correlate with the seriousness of the index offence (although there is no clear reason to expect it to). Correlations between the HCR-20 total score and indices of past crime and violence ranged from .33 to .63. For the H scale, the range was .39 to .68, for the C scale, .30 to .56, and for the R scale, .19 to .45. In general, the correlations were above .40, and many were above .50.

This study is consistent with findings from Canadian violent offenders (see Douglas & Webster, 1999, above) in terms of the relationship with violence. Although the study included post-dictive analyses (as did Douglas & Webster, 1999), the effects are generally large, often exceeding .50, and ranging to .68. The findings support the concurrent validity of the HCR-20, and support the effort of doing larger scale research on the HCR-20 in American criminal offenders.


Summary

This study investigated criminal histories and risk for violence among hate crime offenders. All hate crime offences (N = 814) reported during 1999 to a large metropolitan area in the U.S. were reviewed. Of these, demographic information was available for 581 (71%). Of these 581 cases, 204 resulted in the apprehension of a suspect of the hate offence. These 204 cases comprised the current sample. The police classified an offence as a hate crime only after a multi-step evaluation process.

Demographic details were gathered from the crime report. Criminal histories were recorded from state and federal databases for all crime activity prior to commission of the hate crime. The criminal history was used to score the Cormier-Lang Crime Index (Quinsey, Harris, Rice, & Cormier, 1998), which was used to quantify severity of criminal history. A content analysis of the criminal records was used to score the HCR-20. The evaluation of the criminal histories and scoring of the HCR-20 were completed by raters blind to details about the index crime (other than knowing that the crime had been classified as being a hate crime).

Participants mostly were male (87%) and White (48%; 26.2% Latino; 15% Black, 7.2% multiracial; 2.6% Asian-Pacific). Participants’ mean age was 32.69 years (SD = 14.04 years, range = 12-81 years). Most (57.6%) of the offences were crimes against the person (24.5% verbal threats of harm to the person; 17.9% property crimes).
Thirty-one participants (16%) were identified from the crime reports as being members of a hate-oriented criminal gang or group.

Mean scores on the HCR-20 scales and their alpha reliability coefficients were as follows: Historical (8.21, SD = 4.74, range = 0-19, α = .94); Clinical (5.39, SD = 2.46, range = 0-10, α = .90); Risk Management (5.53, SD = 2.97, range = 0-10, α = .95). The Cohen kappa for inter-rater agreement on the HCR-20 was .58 (range = .33-.90).

The total number of prior arrests and convictions, respectively, were significantly correlated with scores on all scales: Historical (r = .60, p < .001; r = .67, p < .001); Clinical (r = .53, p < .001; r = .48, p < .001); Risk Management (r = .55, p < .001; r = .62, p < .001). Scores on all HCR-20 scales were significantly correlated with severity of prior violent and non-violent crimes, respectively: Historical scale (r = .64, p < .001; r = .66, p < .001); Clinical scale (r = .49, p < .001; r = .57, p < .001); Risk Management scale (r = .56, p < .001; r = .63, p < .001).

The bias intent classified in the crime reports comprised 116 (54.4%) due to racial or ethnic bias, 56 (26.5%) due to the victim’s sexual orientation, 25 (11.3%) due to religious bias, and 6 (2.9%) due to gender bias. ANOVA results indicated no statistically significant differences in mean scores on HCR-20 indices as a function of bias intent.

Finally, comparisons were made between bias offenders who were classified on the crime reports as having been a member of a hate-oriented group or racially motivated criminal gang and the other offenders in the sample. Relative to the other offenders, hate gang members had higher scores on the HCR-20 Historical (t = 3.41, p < .001), Clinical (t = 2.01, p < .01), and Risk Management (t = 4.91, p < .001) scales.

Findings were discussed in terms of their implications for clinical assessment and intervention.

Project and Scholarly Works


Summary

This study evaluated the PCL-R, HCR-20, and VRAG for the prediction of recidivism in a sample of 80 German female offenders who had been referred for forensic-psychiatric evaluation prior to sentencing. The average age of the sample was 36 years. Index offenses were predominately fraud and theft, and 27% were violent index offenses. At the time of the assessment, 13% met the criteria for a personality disorder, 25% for substance abuse or addiction, 9% for schizophrenic disorders, 16% for affective disorders and 19% for other psychiatric disorders. The outcome variable (recidivism vs. non-recidivism) for each participant was obtained from official criminal records. The mean time at risk was 8 years (SD = 5 years). During this time, 31% of the female offenders were reconvicted, 5% for violent (i.e. threat of violence, bodily harm, or sexual offenses) and 26% for non-violent re-offenses (i.e. theft or fraud, or driving without permission). Because only four women committed violent re-offenses during the follow-up period, the authors combined violent and non-violent re-offenses into one general recidivism category.

Mean scores were 13.74 (SD = 6.93), 5.6 (SD = 3.18), 3.44 (SD = 2.30) and 4.73 (SD = 2.63) on the Total, H, C and R scales of the HCR-20 respectively. Mean total scores were 11.99 (SD = 7.95) and 2.00 (SD = 8.49) on the PCL-R and VRAG, respectively. Mean scores were also reported separately for recidivists and non-recidivists. PCL-R scores were examined using the two, three and four factor solutions proposed by Hare (1990), Cooke & Michie (2001) and Hare (2003), respectively. There were significant differences between recidivists and non-recidivists groups on the PCL-R’s antisocial lifestyle factor, F (1, 79) = 4.55, p < .05, lifestyle factor F (1,79) = 5.55, p < .05, and Cooke’s impulsive and irresponsible behavioral style factor, F(1,79) = 5.55, p < .05. There were also differences between groups on VRAG scores, F (1, 79) = 6.35, p < .05, but not on HCR-20 total or subscale scores.

Receiver operating characteristic analyses revealed that the PCL-R total score was significantly predictive for general recidivism (AUC = 0.66; p < 0.05). On the factor level, the antisocial lifestyle factor showed a tendency for a predictive value (AUC = 0.64; p = 0.05), Cooke’s impulsive and irresponsible behavioral style factor showed a significant predictive result (AUC = 0.65; p < 0.05), as did Hare’s lifestyle factor (AUC = 0.65; p < 0.05). No other factors were significant. The VRAG also significantly predicted general recidivism (AUC = 0.72; p < 0.05), whereas the HCR-20 (AUC = 0.59) and its subscales did not show predictive validity (AUCs were 0.61, 0.56 and 0.56 for H, C and R scales, respectively, all ps > .05).

These results appear to provide the first evidence that the PCL-R total score and the antisocial lifestyle factor are predictive for general female recidivism, as has been shown consistently for male recidivists. The HCR-20 however did not significantly predict recidivism. This was inconsistent with most research that has shown a predictive effect of the HCR-20 in female samples. However, because the HCR-20...
is intended to predict violent recidivism, a failure to predict general recidivism does not directly address the validity of the HCR-20.


Summary
In this study, 86 female offenders between 1993 and 2007 were assessed using the PCL-R, the HCR-20, the VRAG, and the ILRV. Approximately 26% of the sample had no psychiatric diagnosis, 24.4% had a substance abuse disorder, 15.1% had a mood disorder and 11.6% had schizophrenia. The participants were at risk for an average of 8 years (SD = 6 years, range 0-18). The results indicated that 34% committed a new offense, 29% of which were nonviolent offences and 5% were violent offences. Only the AUC for the R subscale of the HCR-20 was significant (.79). The AUCs for recidivism for the HCR-20, H and C subscale ranged from .72 to .74 but the confidence intervals were very large, in one case ranging from .37 to 1.00. In addition, the PCL-R total score nor the Hare or Cooke Factors were significant. In contrast, the ILRV D scale was significantly related to the outcome (AUC=.81) and the VRAG was a significant predictor (AUC = .86).

Project and Scholarly Work

Summary
This retrospective cohort study tested the validity of the DUNDRUM-I security triage scale, a structured professional judgment instrument developed to assist placement of patients to appropriate levels of therapeutic security. The DUNDRUM-I consists of 11 items which are rated on a 5-point Likert-scale, from 0 (follow-up with as outpatient) to 4 (allocate to high security). For cross-validation the DUNDRUM-I was compared with the HCR-20 H and C items, coded on a sub-sample of cases (32 individuals), by trained clinicians blind to the outcome. The authors hypothesized that the DUNDRUM-I would correlate weakly or not at all with the HCR-20, which was not designed to take into account the complexity of treatment need.

A total of 316 individuals recruited from facilities in Ireland were fully assessed and rated with the DUNDRUM-I. Internal consistency in the sample was good (α = .95). Of the 316 individuals, 246 were recruited from a busy remand committal prison and were assessed over a 3 month period, and 100 were admitted to a mental health service from the same prison and assessed over a two year period. The dependent variable in the study was transfer to a less secure unit.

For the 246 persons assessed over a three month period, 159 were discharged to the prison GP for follow-up, 57 were followed in the psychiatric in-reach clinic, and 30 were admitted to a psychiatric hospital. The total score on the DUNDRUM-I significantly differed for the three groups. For the 100 persons assessed over a two year period, 53 were transferred to open wards or locked low secure intensive care units and 47 were transferred to medium and high secure units. Again, scores on the DUNDRUM-I differed significantly according to the level of security to which the person was admitted. For both groups, receiver operating characteristics distinguishing between the levels of therapeutic security to which persons were placed was good (AUCs ranged from .81 - .98).

For cross-validation the HCR-20 H and C items were assessed on a sub-sample of 32 individuals. The DUNDRUM-I security triage scale did not significantly correlate with HCR-20 historical items (r = 0.33) nor did it correlate with the HCR-20 clinical items (r = 0.17), indicating that the DUNDRUM-I measured something other than the risk for violence. However, the former correlation is moderate in size and hence its nonsignificance is attributable to low power.

Project and Scholarly Works

Abridged abstract (English translation of the study not available):

The present study sought to describe the risk of violence of filicidal women in an Argentinean prison and determine the
validity of risk assessment instruments for predicting institutional violence within this population. A total of 47 women were assessed using the HCR-20 and PCL-R. Institutional violence was assessed prior to discharge, transfer or the end of the study period, resulting in an average follow up length of 237 days. Numerous women were found to exhibit severe mental health issues (19.5%) or substance use problems (8.5%). The median score on the H scale was 6.8 and the median score on the C scale was 3.7. The median score for the total PLC-R scale was 12.7, with a Factor 1 median score of 6.2 and a Factor 2 median score of 4.9. Overall, the sample was considered to present with low risk for institutional maladaptive behaviours. Of the total sample, 11% in engaged in some type of institutional maladaptive behaviour. The predictive validity of the risk assessment instruments did not reach statistical significance.


Abridged abstract (English translation of the study not available):

Different dynamic factors and violent recidivism were assessed in 25 prisoners on parole from the Province of Buenos Aires Penitentiary System, including the Risk Management scale of the Argentinean version of the HCR-20. The prevalence of risk factors was linked to drug abuse and socioeconomic deprivation. Exposure to destabilizers was the factor associated most strongly with violent recidivism.

Project and Scholarly Work


Summary

The authors investigated prospectively the predictive validity of the HCR-20 (Total, H-scale, and C-scale), PCL-R, Beck Hopelessness Scale (BHS; Beck, Weissman, Lester, & Trexler, 1974), and Brief Psychiatric Rating Scale (BPRS; Overall & Gorham, 1962) with respect to institutional misconduct and self-harm/suicide. The R-scale was not scored because evaluations were completed within 2 weeks of admission at which time participants’ social situations at discharge were difficult to project. The sample consisted of 34 mentally disordered offenders admitted to one of two medium-secure hospital units in the U.K. All participants had committed a serious offense (e.g., manslaughter, sexual offense, physical assault, arson). Almost half were diagnosed with paranoid schizophrenia (44.1%; 23.5% depression; 14.7% personality disorder; 17.6% ‘other’, such as bipolar disorder, organic disorder, neurosis). At admission, the mean age was 33 years ($SD = 11.9$). Participants primarily were male (76.5%) and White (85.3%).

Incidents relating to four outcome criteria were measured: (1) verbal aggression, (2) aggression to property, (3) physical aggression to staff or other patients, and (4) self-harm or suicide. All incidents were assessed using a structured record form developed for this study (the Aggression Vulnerability Scale; AVS). The authors reported good intraclass correlations for the AVS for all four outcome criteria in a pilot study. AVS ratings were completed weekly for 3 months (or until the participant was transferred off the unit) using nursing records, incident report forms, and weekly interviews with the primary nurses.

The mean scores for the measures were as follows. H-scale ($10.2; SD = 3.4$; range 3-18); C-scale ($5.4; SD = 2.6$; range 1-10); PCL-R total ($10.4; SD = 1.5$; range 0-33); Factor 1 ($3.9; SD = 4.3$); Factor 2 ($5.5; SD = 4.2$); BHS ($8.6; SD = 6.4$; range 1-19); BPRS ($42.0; SD = 2.7$; range 17-74). Correlations between the measures tended to be high and significant. Notably, there was an especially strong relationship between the BPRS and the HCR-20 total score (.63) and C-scale (.71).

The authors used correlations, ROC analysis, odds ratios, and the Mann-Whitney U Test (a nonparametric comparison of means for those scoring higher versus those scoring lower than the median value) to examine predictive validity for the HCR-20, PCL-R, BHS, and BPRS with respect to each of the four dependent measures. For verbal aggression and HC total score: $r = .53, p < .001$; AUC = .79, $p < .001$; OR = 2.55; Mdn split $p < .01$. For verbal aggression and H-scale: $r = .43, p < .01$; AUC = .73, $p < .05$; OR = 2.21; Mdn split $p < .01$. For verbal aggression and C-scale: $r = .49, p < .01$; AUC = .74, $p < .01$; OR = 2.27; Mdn split $p < .05$. For violence to property and HC total score: $r = .56, p < .001$; AUC = .83, $p < .001$; OR = 8.85; Mdn split $p < .01$. For violence to property and H-scale: $r = .54, p < .001$; AUC = .82, $p < .001$; OR = 8.45; Mdn split $p < .01$. For violence to property and C-scale: $r = .49, p < .001$; AUC = .77, $p < .001$; OR = 3.85; Mdn split $p = .09$. For physical aggression and HC total score: $r = .53, p < .001$; AUC = .81, $p < .001$; OR =
8.25; Mdn split $p = < .01$. For physical aggression and H-scale: $r = .43$, $p < .01$; AUC = .77, $p < .001$; OR = 7.46; Mdn split $p = < .05$. For physical aggression and C-scale: $r = .49$, $p < .01$; AUC = .79, $p < .001$; OR = 7.42; Mdn split $p = < .05$. The HCR-20 was not predictive of self-harm, which is not surprising given that it was developed to predict risk of violence to others.

Values for the predictive indices yielded for the HCR-20 were larger than those for the PCL-R. The PCL-R total score was, however, a useful predictor for violence to property and physical violence ($r = .35 - .38$; AUCs = .70-.76; ORs = 1.88 – 2.84). The PCL-R was not a significant predictor of verbal aggression or self-harm.

Project and Scholarly Work


Summary

The authors examined rulings and expert testimony given in 304 cases which led to Preventive Detention between 1991 and 2001 in 4 states of Germany. Of the 304 cases, 224 ruling were available to be analysed and 227 expert testimonies were analyzed (3 of the cases were seen by 2 experts). Eighty-seven percent of the cases were violent offences (17% murder/homicide cases and 30% Robbény) and 50% were sexual offences. Juvenile delinquents composed the majority of cases (84%), 45.1% had violated parole and 15.5% were delinquent in prison. Preventive Detention is intended for “social disintegrated recurrent offenders with noticeable personality problems or personality disorders.” About half of the sample had a mental disorder, mostly a Cluster B personality disorder and 26% had a substance disorder. In testimonies that did not provide a diagnosis (96 cases) 72 cases included descriptions of personality traits. Generally 66% had noticeable personality problems.

The content analysis focused on whether data relevant for the HCR-20, PCL-R, SVR-20, and Static-99 was found in the expert’s testimony. The testimony did not include a standardized assessment of risk factors (HCR used in 1 case) but the average testimony contained information about 9.9 PCL-R and 11.6 HCR items. Items that were seldom mentioned were Psychopathy, Need for stimulation, Conning/manipulative, Superficial charm, Pathological lying, Lack of personal support, Parasitic lifestyle, Stress, Unresponsive to treatment and Exposure to destabilizers. For the SVR-20, items that were mentioned less than 20% of the time were Negative attitude toward intervention, Escalation in frequency or severity of sex offences, Lacks realistic plans, Suicidal/homicidal ideation, Attitudes that support or condone sex offences, and Psychopathy. In the Static-99, the type of victim and convictions for non-contact sex offences were mentioned less than 20% of the time.

HCR-20 Review and Annotated Bibliography

Project and Scholarly Work


Summary

The present study outlines the development, implementation, and outcomes of a prison diversion treatment program in Central California. The resulting 12-week cognitive-behavioral program was designed in accordance with empirical literature. Twenty-three offenders serving suspended prison sentences completed the program, and statistical analyses of pre-post differences revealed significant improvements in graduates’ ratings of quality of life.

Clinician-rated measures of treatment outcomes were established via comparisons of before and after scores on the dynamic factors of the HCR-20: clinical items (lack of insight, negative attitudes, active symptoms of mental illness, impulsivity, unresponsive to treatment) and risk management items (plans lack feasibility, exposure to destabilizers, lack of personal support, noncompliance with remediation attempts, stress). Significant improvements in clinical items regarding negative attitudes and impulsivity were found, indicating that clinicians noticed an observable improvement in the outlook and behaviors of the participants. These findings are also in alignment with the self-ratings of participants on the program survey wherein participants noted an increase in confidence, optimism, and self-efficacy. Overall, participants reported experiencing improvements in self-awareness, self-esteem, decision making, and goal setting. Preliminary 6-month findings of the initial 13 participants have yielded a 0% recidivism rate, with 100% of the program graduates maintaining gainful employment, pursuing educational goals, or both. The author concludes by discussing considerations for subsequent program development are discussed and study limitations.
Project and Scholarly Work


Summary

The current retrospective study investigated two risk assessment instruments (Salient Factor Score and the CR subscales of the HCR-20) in 64 adult male and female inmates. The sample was incarcerated in San Francisco County Jails, diagnosed with a serious Axis I disorder, and enrolled in Behavioral Health Court (BHC). Participants were randomly chosen from the BHC. The Salient Factor Score (SFS; static risk factors) and the C and R subscales of the HCR-20 (dynamic risk factors) were coded based on charts from the jail. The study evaluated the predictive validity of the two risk assessment instruments for general and violent recidivism in the community over a 2 year period. The outcome was re-arrest (for a felony or serious misdemeanor) which was coded from San Francisco county arrest records. General recidivism was defined as crimes that did not involve threatening or assaultive behaviour against another person or animal. Violent recidivism was defined as behaviours that were aggressive in nature (e.g., possession of a weapon, physical assault, Robbery, aggressive interpersonal interactions).

The sample was 68.8% male with an age range of 20 to 69 (32.8% between 30-39 years of age). The ethnicity of the sample was 45.3% Black-Americans, 23.4% Caucasians, 12.5% Asian/Pacific Islanders, and 10.9% Latinos. The majority of the sample had prior arrests (n = 58) and a history of violence (n = 50). Schizophrenia was the most prevalent Axis I disorder (42.2%), followed by schizoaffective disorder (23.4%), psychotic disorder not otherwise specified (12.5%), major depressive disorder (6.3%), bipolar I (4.7%), and mental retardation (3.1%).

Inter-rater reliability on the HCR-20 was assessed with two raters coding 21 randomly chosen participants. The IRR of the CR-10 composite was good with interclass correlations between .77 and 1.00, with the exception of Item 3 on the Clinical Scale (active symptoms of major mental illness). The mean score on the SFS was 6.78 (SD = 2.56), which falls within the “good prognosis” category. The mean of the CR-10 was 10.30 (SD = 3.37). Overall, the R subscale was the only predictor of general recidivism (with the SFS and C subscale in the model) and the C subscale was the only predictor of violent recidivism (with the SFS and R subscale in the model). For general recidivism, the instruments showed moderate predictive ability (Cr-10 AUC = .67, R subscale AUC = .73; SFS AUC = .67. For violent recidivism, the CR-10 produced an AUC of .65 and the R subscale produced an AUC of .68. The SFS was not a significant predictor.

Project and Scholarly Works


Summary

Kroner and Mills (2001) completed a prospective study of institutional misconduct among offenders. At intake, they completed the HCR-20, along with other instruments, on a sample of 97 consecutively admitted Canadian federal offenders. In terms of interrater reliability, ICC1 was reported to be .85 for the total score. The HCR-20 correlated .32 with minor misconducts, and .11 with major misconducts, and was not significantly different than the PCL-R, the VRAG, or the LSI-R. For post-release analyses, the HCR-20 correlated at .28, .16, .21, and .39 with total convictions, violent convictions, nonviolent convictions, and revocations, respectively. Again, there were no significant differences between measures.

It should be pointed out that, although there were no significant differences between measures, that the coding procedure was not optimal for community violence. The instruments were coded at admission, and this score was used to predict violence after release from the institution, some years later. This has particularly strong implications for measures that will change over time, such as the C and R scale and the LSI-R. Further, institutional outcome criteria (misconduct) included mostly non-violent indices, such as improper dress, disrespect, noncompliance with directions, drug use, and refusing urinalysis.


Summary

This study compared the validity of the individual facet scores of the PCL-R in predicting violent and nonviolent recidivism. Additionally the study assessed whether Facets 1, 2, and 3 provided any incremental validity to Facet 4 in the prediction of recidivism and vice-versa. Finally, this study examined whether incorporating Facet 4 as opposed...
to the total PCL-R score yielded any meaningful increase or decrease in the predictive validity of either the HCR-20 or the VRAG.

The study sample consisted of 248 participants drawn from a population of federally incarcerated men with sentences ranging from 2 years to life. The mean age of the sample was 29.6 (SD = 8.0). A majority of the sample were Caucasian (71.7%). The HCR-20, PCL-R and VRAG were administered by a trained clinician as part of the offenders’ psychological risk assessment. Recidivism was coded from both the offenders’ correctional files and official police records. Only offences committed post-testing and within the community which resulted in a conviction were recorded; those committed while the offender was in custody or prior to testing were not recorded. The average follow-up time was 11.1 years. Of the sample, 44.6% and 69.2% committed violent and non-violent offenses during the follow-up period, respectively.

For any violence, AUCs were .56, .54, .59, .63, .79, and .68 for Facet 1, Facet 2, Facet 3, Facet 4, HCR-20 and VRAG scores, respectively. For any non-violent recidivism, AUCs were .61, .58, .64, .68, .69, and .74 for Facet 1, Facet 2, Facet 3, Facet 4, HCR-20 and VRAG scores, respectively. Of the four Facets, only Facet 3 (Lifestyle) and Facet 4 (Antisocial) were consistent predictors of recidivism. However, with the exception of the Affective and Antisocial comparisons, the observed differences were not significant. Facet 4 was found to add incremental validity to the prediction of recidivism relative to the remaining three Facets, yet Facets 1, 2, and 3 did not add incremental validity relative to Facet 4. Pairwise ROC Curve Comparisons for any recidivism and any non-violence recidivism indicate that the predictive accuracy of neither the VRAG nor the HCR-20 was dependent upon the PCL-R total score when incorporating only the Antisocial Facet score.


Summary

Authors reported correlations that were not published in their original report, describing the relationship between risk measures (HCR-20 and VRAG) and violent reoffending over a longer period than published in the article.

Significant correlations between the HCR-20 items and violent behaviour are as follows. Total score (r = .37; p < .01), H2 (r = .28; p < .05), H3 (r = .28; p < .05), H4 (r = .23; p < .05), H8 (r = .22; p < .05), C1 (r = .27; p < .05), C2 (r = .34; p < .01), C4 (r = .28; p < .05), C5 (r = .25; p < .05), R1 (r = .26; p < .05), R2 (r = .26; p < .05), R4 (r = .35; p < .01) and R5 (r = .27; p < .05).

HCR-20 Review and Annotated Bibliography

Significant correlations between the VRAG and violent behaviour are as follows: Total score (r = .28; p < .05), item #2 (r = .29; p < .01), item #6 (r = .28; p < .05) and item #11 (r = .27; p < .05).


Summary

The current study examined the ability of individual items on the HCR-20 and VRAG to discriminate between violent recidivists and nonrecidivists within an offender population. Participants were 83 volunteers drawn from a population of incarcerated men in federal custody in Canada (sentenced to 2 years or more). The ages of the participants ranged from 18 to 55 years, with a mean age of 27.9 years (SD = 8.1). The majority of the sample was Caucasian (n = 67) followed by African American (n = 7), Native North American (n = 5), and Asian (n = 4). Participant’s most serious index offenses were assaultive (n = 51), Robbery (n = 18), criminal negligence/driving (n = 7), arson (n = 6), and drug related (n = 1). The mean number of convictions and incarcerations for these participants was 12.9 (SD = 13.5) and 4.6 (SD = 5.3), respectively.

Participants were followed-up in the community via offender correctional files and official police records. Participants were classified as either violent recidivists or nonrecidivists keeping with the methodology of the original VRAG study. Violent offenses included uttering threats, assault (on someone other than a peace officer), sexual assault, armed Robbery, and Robbery with violence. Withdrawn and dismissed charges were counted as offenses only when there was compelling file information to indicate that the participant did commit the offense. The average follow-up period for all offenders in the sample was 4.6 years (SD = 337.4 days). The number of days to an act of violent recidivism ranged from 16 days to 4.3 years (M = 561.2 days, SD = 464.7 days).

Assessments were conducted within 8 to 12 weeks of arrival for the purpose of identifying risk level and determining management and programming in federal custody. Ratings were made by trained clinicians. Ratings of the HCR-20 and VRAG were based on information from an interview, correctional file review, and police records. The base rate for violent offending was 35% (29 violent recidivists and 54 nonrecidivists). For the violent recidivists, the mean score on the HCR-20 was 18.3 (SD = 8.4, range = 0-37). The mean score for the H subscale was 9.2 (SD = 4.4, 0-18), for the C subscale 4.1 (SD = 2.2, 0-9) and for the R subscale the mean score was 4.9 (SD = 2.7, 0-10). The mean score of the VRAG was 8.43 (SD = 11.1, 12-31).
The HCR-20 (AUC = .72) and all of its subscales (H-10, AUC = .67; C-5 AUC=.75; R-5, AUC=.71) were significantly related to violent recidivism. Similarly, the VRAG was also significantly predictive of violent recidivism (AUC=.67). At the item level, the analysis showed that 15 of 20 HCR-20 items were significantly related to the number of days to a violent offense, whereas only 5 of 12 items of the VRAG were related to the number of days to a violent offense. Among the items of the HCR-20 most strongly related (r = .30 or higher) to time to reoffence were noncompliance (R4), exposure to destabilizers (R2), age at first violent offense (H2), previous violence (H1), and negative attitudes (C2). Only one item of the VRAG was strongly related to days to violent reoffence: failure on prior conditional release (V6; r = .41). The final step of the analyses was to modify both the HCR-20 and VRAG, removing the items that were not related to violent outcome in our sample. The strength of the relationship between the modified versions of the instruments and violent outcome was then analyzed in the same manner as that for the original instruments. The relationship with the outcome improved little over the original scales for both instruments: HCR-20 (AUC = .73); VRAG (AUC = .69).


Summary

The Two-Tiered Violence Risk Estimates (TTV) instrument is a violence risk appraisal instrument designed to support an integrated-actuarial approach to violence risk assessment. The TTV provides both an actuarial risk estimate (ARE) in conjunction with a structured approach to assessing risk management indicators (RMI) that explain or are associated with the underlying risk estimate. ARE scores can range from 0 to 13 and the 13 RMI items are used to identify dynamic risk factors that require intervention, monitoring or management. This study was an initial validation study of the TTV. As part of this study construct validity was examined by comparing the TTV with the HCR-20 and the Lifestyle Criminality Screening Form (LCSF).

The sample used in this study consisted of 78 participants drawn from a population of federally incarcerated men in Canada. The outcome was recidivism, which was coded from correctional files and official police records. The average follow-up period was 12.42 years (SD = 318.80 days). Only offenses committed post assessment and within the community which received a new conviction were recorded. Time-at-risk was calculated for both violent and any recidivism and was calculated as the number of days between the initial date of release from a federal institution and one of three subsequent dates: date of offense/conviction (violent or any), date of death, or date of follow-up. The base rates of violent and any recidivism were 47.4% and 73.1%, respectively. Among the 78 offenders, a total number of 108 (M = 1.38, SD = 1.95) convictions for violent recidivism and 644 (M = 8.26, SD = 10.11) any recidivism were recorded over the follow-up period.

Both the HCR-20 and LCSF were completed by one of three trained psychologists, whereas the TTV was retrospectively scored from file based on the information that was available at the time of the assessment. Interrater reliability for the HCR-20 and LCSF was high (ICC = .85 and .85, respectively). Overall internal consistency of the TTV (α = .77 and .76 for ARE and RMI, respectively), HCR-20 (α = .88, .77, .62 and .80 for Total, H, C, and R scales) and LCSF (α = .74 for total score) was adequate. Mean scores on the instruments were as follows: ARE M = 8.62, SD = 3.14; RMI M = 8.47, SD = 2.75; HCR-20 Total M=19.28, SD = 7.67; Clinical M = 4.32, SD = 2.00; Risk Management M = 5.28, SD = 2.50; and LCSF M = 12.19, SD = 4.33. Both ARE and RMI were highly correlated with the total and subscale scores of the HCR-20 (correlation with ARE r = .73,.77, .49 and .56 for Total, H, C, and R scales, respectively) and LCSF (r = .81 and .71 for ARE and RMI, respectively).

The TTV was found to be a robust predictor of violent recidivism. The RMI produced the largest AUC value with recidivism (AUC = .80), whereas the ARE produced the second largest AUC value (AUC = .80), although neither of these were either meaningfully or significantly different than the AUC of .79 for the HCR-20. Similar trends were evident for the correlation coefficients between the RMI and ARE with time-at-risk until violent recidivism (r = .53 and – 51, respectively) and number of violent convictions (r = .44 and .43, respectively). Both the HCR-20 and LCSF total and subsalce scores displayed moderate to large AUC values with respect to violent recidivism. AUCs were .79, .77, .75, and .69 for total, H, C, and R scales of the HCR-20. The LCSF had an AUC value of .76. With respect to the correlation coefficients, both the HCR -20 (r =.48 and -.43 for violent recidivism and time-at-risk, respectively) and .LC SF (r = .70 and -.50 violent recidivism and time-at risk) were also significant.

AUC values did not significantly differ between the HCR-20 and the LCSF for either violent or any recidivism. Furthermore, no statistically significant differences were noted between the AUC values generated by the ARE and RMI with violent or any recidivism. The authors therefore considered the predictive validity of the total and substance
scores for the TTV, HCR-20 and LCSF to be statistically equivalent. The authors concluded that the TTV was a reliable and valid risk assessment measure that was predictive over the long term.

See Also


Project and Scholarly Works


Summary

The HCR-20 consists of many dynamic risk factors that are assumed to change over time; however, little research has investigated whether in fact these risk factors do change and how changes in these risk factors relate to the individual’s overall risk. The current study aimed to characterize any changes in the HCR-20 scores over time and how these changes may be related to recidivism. Participants consisted of 158 offenders on probation (47.5%) or parole (52.5%). The majority were male (86.7%), under supervision for a drug related offence (41.1%) or property related offence (35.4%), and had a history of violence perpetration (69.6%). Participants were assessed at baseline and then followed for an average of 13 months with reassessment at 6 months and 13 months into the follow-up period.

HCR-20 total scores did not change significantly over the course of the assessments (Baseline $M = 14.73$, $SD = 6.47$; First reassessment $M = 15.02$, $SD = 6.75$; Second reassessment $M = 15.12$, $SD = 6.87$). Very little changes occurred in the final risk judgments over the follow-up period. From the initial assessment to the first reassessment 85.2% of the risk judgments stayed the same; also from the first to second reassessments 85.2% of the risk judgments stayed the same. At the scale level, the only significant differences were found on the R scale between the baseline assessment and first reassessment ($r = -.235$, $p < .05$) and the baseline assessment and second reassessment ($r = -2.78$, $p < .05$), with the scale scores increasing from baseline. At the item level, the only risk factors that differed significantly over the assessments were H10 (prior supervision failure), which increased over the assessments, and R5 (stress), which peaked at the first reassessment and then fell. Several other items differed across the assessments but not to a level that attained statistical significance.

Next, the researchers investigated whether any of these changes contributed to the predictive validity of the scores. The baseline assessment was correlated with general recidivism with coefficients of .53, .43, .46, and .45 for the total score, H, C, and R scales, respectively. The first reassessment was also predictive of general recidivism with coefficients of .55, .43, .54, and .45 for the total score, H, C, and R scales, respectively. Therefore, the general trend indicated that the first reassessment was more predictive of recidivism; however, none of the differences between the correlation coefficients were statistically significant. AUC confirmed the trend with the baseline assessment providing an AUC = .83 and the first reassessment providing an AUC = .85. Moreover, recidivists showed a small increase in mean score from baseline to first reassessment, while the non-recidivists showed no change in mean score over the assessments.

The authors concluded that more changes in HCR-20 scale and total scores were expected over the assessments given the dynamic nature of many of the risk factors. Nevertheless, there was a trend for reassessment to be more accurate at predicting the outcome of interest.


Summary

In the present study the predictive, convergent, and incremental validity of the HCR-20 and the PCL-R were prospectively tested in relation to general, violent, and nonviolent recidivism in a sample of Portuguese offenders. The sample used in this study consisted of 158 probationers and parolees under supervision by the Portuguese Probation/Parole Service units. Selection criteria included serving a community sentence that would last for at least one year and being at least 18 years of age. The mean age of the participants was 35.15. A majority of the sample was male (86.7%), registered at least one event of previous
violence (69.6%) and had committed drug-related offenses (41.1%). Of the sample, 5.1% had a history of a major mental illness and 1.9% was diagnosed with a personality disorder.

Participants were assessed using the HCR-20 and PCL-R and prospectively followed up for an average period of 13 months ($M = 12.82, SD = 1.58$). Data for the assessments were collected by the study’s first author through individual interviews, file records reviews and close interactions with the probation/parole officers. Participant self-report and all official information that came to the knowledge of the Probation Service was used to code recidivism. In terms of outcome, 35.4% showed at least one occurrence indicative of general recidivism. Non-violent recidivism, which included technical violations and nonviolent criminal behavior, was detected in 24.1% of participants. Violent behavior, defined by the commission of a violent crime or by the involvement in mutual physical aggression incidents, was present in 17.7% of participants.

The mean total score of the HCR-20 was 15.34 ($SD = 6.64$). Mean scores were 7.52 ($SD = 3.70$), 3.61 ($SD = 2.00$), and 4.27 ($SD = 2.11$) on the historical, clinical and risk management scales, respectively. The mean total score on the PCL-R was 13.96 ($SD = 6.75$). Mean scores were also reported for each of the factors and facets of the PCL-R. The authors analyzed the correlations between HCR-20 scores and PCL-R scores, with Item H7 (Psychopathy) of the HCR-20 removed prior to analysis. All correlations were statistically significant. The historical scale of the HCR-20 and Factor 2 of the PCL-R showed the highest correlation ($r = .79, p < .01$) followed by the correlation between HCR-20 total and PCL-R total scores ($r = .75, p < .01$).

ROC analyses revealed that both the HCR-20 and the PCL-R showed good predictive validity, however the HCR-20 tended to outperform the PCL-R for general, violent, and nonviolent recidivism. AUC values for the HCR-20 total scores were .84, .82 and .81 for general, nonviolent, and violent recidivism, respectively, while the AUC values for the PCL-R total scores were .81, .80 and .77 for general, violent and non-violent recidivism, respectively. Considering each subscale of the HCR-20, general recidivism and nonviolent recidivism were better predicted by risk management items (AUCs were .80). Violent recidivism showed the highest AUC with historical items (AUC = .83). SPJ reached consistently higher predictive validity coefficients (AUC = .83) rather than the use of the numeric scores of the HCR-20 (AUC = .81) for violent recidivism. For general and nonviolent recidivism, however, the SPJ had AUCs of .83 and .78 respectively, which were lower than the validity coefficients using the numeric scores. AUC values for all total and subscale scores of the HCR-20 and PCL-R were significant. Point biserial correlations ($r_{pb}$) for each of the measures and outcome were also reported. SPJ was the best predictor for general recidivism ($r_{pb} = .59$). Correlations for nonviolent recidivism were also slightly higher for total scores ($r_{pb} = .48$) in comparison with the SPJ summary ratings ($r_{pb} = .45$). All correlations were significant.

To examine incremental validity of the HCR-20 in comparison to the PCL-R the authors performed a series of logistic regression analyses for each recidivism outcome. First, characterization variables (i.e., sex, age, nationality, supervisory status), then the PCL-R and the HCR-20 total scores were entered into the model. For each outcome, when the HCR-20 was entered into the model, the PCL-R lost significant contribution in the model. Further analyses were performed reversing the order of entry of the PCL-R and the HCR-20. Results confirmed previous findings in all three outcomes: the PCL-R was not a significant predictor when added to the model. The authors concluded that the HCR-20 adds significant incremental validity to predictions made solely with the PCL-R.

In addition, the performance of the structured professional judgment (SPJ) was compared to the numerical use of the HCR-20. Incremental validity analyses of SPJ summary risk ratings were tested in relation to the HCR-20 total score. HCR-20 total score, then HCR-20 total score and SPJ summary risk ratings. The authors found that SPJ ratings added incrementally to the prediction of violent behavior. General recidivism was also better predicted when adding the SPJ. The prediction of nonviolent recidivism also improved with the inclusion of SPJ but the SPJ variable did not arise as a significant predictor.

The authors concluded that the HCR-20 and the PCL-R can be used with Portuguese offenders with valid and efficient results. High coefficients for the prediction of general and nonviolent recidivism showed that the HCR-20 and the PCL-R can be valid risk measures for outcome other than violence and that they can be useful in nonclinical settings. Results also provided support for the validity and utility of the SPJ method, especially for the assessment of violent outcomes.

See Also


Project and Scholarly Work

Abridged abstract (English translation of the study not available):

The convergent validity between the PCL-R, HCR-20, and DSM-IV Axis I and II diagnoses was studied in a sample of 60 male prisoners in France. Convergence between the HCR-20 and the PCL-R was high. The Historical scale correlated more strongly with PCL-R Factor 2, whereas the Clinical scale correlated more strongly with PCL-R Factor 1. The diagnoses of antisocial, borderline and narcissistic personalities, as well as drug dependence, were associated with HCR-20 total score. "Suicidal Risk" and "Generalized Anxiety" correlate negatively with the PCL-R Factor 2, whereas the "Factor 1" scores. HCR-20 total scores, and especially scores on the Historical scale, were associated with having committed Robbé and assault. PCL-R total score was associated with commission of Robbé. Neither measure correlated positively with having committed homicide or sexual offences.

Project and Scholarly Work


Summary

This study collected HCR-20 and PCL-R data from 68 offenders in a Belgian high security prison. The mean follow-up period, available for 38 offenders, was 1010 (SD = 894) days. Interrater reliability, based on a subsample of 10 offenders, was .85 (Pearson r) for the HCR-20 Total Scale Score. The AUCs between HCR-20 indices and violent recidivism were as follows: HCR-20 Total Scale (.76); Historical Scale (.77); Clinical Scale (.74); Risk Management (.71). PCL-R AUCs were .82 (Full Scale), .77 (Factor 1), and .75 (Factor 2). Based on a subset of 20 offenders, correlations between the HCR-20 indices and the Buss and Perry Aggression Questionnaire ranged from .29 to .57, and with the Heilbrun Index of Dangerosity in the Community, from .32 to .37.

HCR-20 Review and Annotated Bibliography


Abridged abstract (English translation of the study not available):

Background: The HCR-20 is a risk assessment instrument developed in Canada whose validity has been shown in several studies of patients in forensic settings. However, only a few studies on unselected samples exist, and they show varying results. The aim of the present study was to conduct a first-time validation of the HCR-20 on an unselected sample of offenders released from prison in Switzerland. Method: The HCR-20 and the PCL-R were scored for 107 released sex and violent offenders. Legal probation was examined using excerpts from the criminal records drawn after a follow-up period of 7 years. ROC analyses were computed to examine the predictive validity of the sum scores of the two instruments. Results: The HCR-20 was significantly associated with recidivism and showed an AUC = 0.76 for general recidivism and an AUC = 0.70 for sex or violent recidivism. The difference in predictive validity of the HCR-20 and the PCL-R was not significant and the two instruments correlated with r = 0.69. Discussion: It was also possible to replicate findings of other studies on the predictive validity of the HCR-20 in Switzerland. However, in the present study the HCR-20 was of no additional explanatory value compared to the PCL-R.

Project and Scholarly Work


Abridged abstract (English translation of the study not available):

This study presents data on German offenders from a larger initiative undertaken in Germany, United Kingdom, The Netherlands, and Norway to develop a unified approach to the assessment of social risk and related behaviours in offender groups. The PCL-R, HCR-20; SCL-90-R; BDHI-D, and Behavioural Status Index (BEST-Index) were administered to 231 mentally ill offenders. The present study reported results from a sub-sample of 89 German offenders. The BEST-Index shows convergent validity with
Project and Scholarly Work


Summary

This study retrospectively examined the predictive validity of an Intellectual Disability (ID) Supplement for the HCR-20 (not developed by the HCR-20 authors) for in a sample of 59 individuals with ID who had committed a violent offense. The participant sample was derived from a post-custodial release program for offenders with ID in New South Wales and was primarily male (93%). The majority of clients (68%) had mild ID, 22% had moderate ID, 9% had mild/moderate ID and 2% were borderline ID with strong adaptive functioning deficits. Forty-four percent of these individuals were imprisoned following index offenses and had a mean age of 24.68 years (*SD* = 8.06) at the time of the offense.

The HCR-20, HCR-20 with ID Supplement, and VRAG were coded from a comprehensive review of client case files by the primary author, who was blind to recidivism outcomes. Rating of the psychopathy item of the HCR-20 and HCR-20 ID supplement were supervised by a senior psychologist certified in scoring the measure. To assess inter-rater reliability, two qualified raters scored a random sample of seven cases. Inter-rater reliability was moderate for total scores on the HCR-20, HCR-20 with ID Supplement, and the VRAG; ICC = .67, .65, and .66, respectively. Recidivism, as indicated by any charge or conviction for new offenses, was collected from criminal justice program files. Violence was defined as any actual, attempted, or threatened harm to a person or persons. In total, 78% of the sample violently reoffended and 92% committed a general re-offense.

Total scores on HCR-20 and HCR-20 with ID Supplement were found to be highly correlated (*r* = .98). Total scores on the HCR-20 and HCR-20 with ID Supplement were also highly correlated with those on the VRAG (*r* = .65 and .66, respectively, both *p* < .001). Both violent and general recidivists obtained significantly higher scores on all scales than did non-recidivists irrespective of type of offense. Violent recidivists had a mean score of 28.02 (*SD* = 4.12) on the HCR-20, a mean score of 27.67 (*SD* = 3.96) on the HCR-20 with ID Supplement and a mean score of 14.32 (*SD* = 6.53) on the VRAG. General recidivists had a mean score of 27.76 (*SD* = 4.21) on the HCR-20, a mean score of 27.46 (*SD* = 3.05) on the HCR-20 with ID Supplement and a mean score of 13.79 (*SD* = 6.66) on the VRAG. For the sample as a whole, the means were: HCR-20 total *M* = 26.93 (*SD* = 5.13); H subscale *M* = 14.69 (*SD* = 3.01); C subscale *M* = 6.15 (*SD* = 1.74); R scale *M* = 6.19 (*SD* = 1.75); HCR-20 total with ID Supplement *M* = 26.59, (*SD* = 5.05); and VRAG *M* = 12.61 (*SD* = 7.84).

The predictive accuracy of the HCR-20 with ID Supplement was compared to that of the HCR-20 and the VRAG. All risk assessment scales, except the Clinical scales of the HCR-20 and the ID Supplement significantly predicted violent and general recidivism as indicated by AUC values and Pearson point-biserial correlations. Accuracy of prediction appeared slightly higher for the HCR-20 with ID Supplement than the HCR-20 when considering AUCs, however because confidence intervals for the instruments overlapped the significance of differences between AUCs could not be determined.

With regards to violent recidivism, AUCs were .80, .81, .80, .82, and .79 (all *p* < .01) for HCR-20 total scores, HCR-20 final risk categories, HCR-20 with ID Supplement total scores, HCR-20 with ID Supplement final risk categories and VRAG scores, respectively. The probability of correct classification for violent recidivism was 80% for the HCR-20 and the HCR-20 with ID Supplement and 79% for the VRAG. With regards to general recidivism, AUCs were .94, .88, .97, .88 and .92 (all *p* < .01) for HCR-20 total scores, HCR-20 final risk categories, HCR-20 with ID Supplement total scores, HCR-20 with ID Supplement final risk categories and VRAG scores, respectively. The probability of correct classification of general recidivism was 90% for the HCR-20, 88% for the HCR-20 with ID Supplement and 91% for the VRAG.

The authors concluded that the ID Supplement can be used as a valid guideline to accompany the original HCR-20 in assessment of risk in ID offenders. However, it was not more strongly related to violence than is the HCR-20 without the use of the supplement.

See Also

Project and Scholarly Work


Summary

The potential usefulness of the PCL-R and HCR-20 in determining level of risk for violent behaviour and other forms of criminality was investigated. Participants were part of a larger study that examined DSM-IV personality disorders using the SCID-II. In conducting PCL-R and HCR-20 interviews, all 261 inmates who had completed the SCID-II interviews approximately 12 months earlier and who were still housed at the maximum-security prison were approached and invited to participate in this subsequent stage of data collection. The final sample comprised 132 women. Sixty percent of the women were under the age of 32 years and 65% were of minority status. Seventy-seven percent of the sample was serving sentences of greater than 5 years and 83% had criminal histories containing at least one conviction for a violent crime.

Each inmate’s file was reviewed by six coders who summarized information about the inmate’s family history, psychiatric history, employment history, and criminal record. The PCL-R and HCR-20 coders reviewed these summary files before they conducted their assessment interviews. Scores on three HCR-20 items were obtained from alternative sources that were thought to be superior to those obtained through a clinical interview: H5 was coded from data obtained for the administration of the Diagnostic Interview Schedule (DIS-IV) for the Alcohol and Substance Abuse module; H9 was scored based upon data obtained in the SCID-II interview; C1 was coded as 0 if inmates received a total Barratt Impulsivity Scale score below 40, 1 if they scored between 40 and 79, and 2 if scoring above 80.

Reliability coding of 28 cases yielded the following intra-class correlation coefficients (ICCs): HCR-20 Total (0.94); Historical scale (0.92); Risk Management scale (0.60); Clinical scale (0.76); PCL-R Total (.95); PCL-R Factor 1 (.88); PCL-R Factor 2 (.99).

Information for both the instant offence and prior offences was obtained from inmates’ prison files. Violent crimes were defined as murder, assault, and battery. Potentially violent crimes included Robbery, kidnapping, and arson. Crimes against persons were defined as negligent homicide, contributing to the delinquency of a minor, hit and run, coercion, unlawful restraint, harassment, criminal possession of weapon, menacing, and reckless endangerment. Property crimes included breaking and entering, tampering, trespassing, larceny, auto theft, shoplifting, possession of stolen property, forgery, fraud, uttering, bribery, and conspiracy. Minor crimes were considered to include parole and probation violations, driving while intoxicated, public drunkenness, failure to appear, gambling, resisting arrest, loitering, public lewdness, traffic infractions, and prostitution. General categories of sex crimes (rape, sexual assault) and drug crimes (possession) also were coded. An overall category of total violent crimes subsumed the violent, potentially violent, crimes against persons, and sex categories, and an overall category of total non-violent crimes subsumed the property, drugs, and minor crime categories.

Descriptive characteristics and inter-rater reliability for the HCR-20 indices were as follows: Total (M = 20.36, SD = 6.87, SEM = 0.60, range 2-35, ICC = .94); Historical scale (M = 11.21, SD = 3.62, SEM = 0.32, range 2-19, ICC = .92); Clinical scale (M = 3.53, SD = 1.90, SEM = 0.17, range 0-9, ICC = .60); Risk Management scale (M = 5.61, SD = 2.52, SEM = 0.22, range 0-10, ICC = .76).

Descriptive characteristics and inter-rater reliability for the PCL-R indices were as follows: Total (M = 22.80, SD = 6.98, SEM = 0.61, range 3-36, ICC = .95); Hare Factor 1 (M = 9.31, SD = 3.82, SEM = 0.33, range 1-16, ICC = .88); Hare Factor 2 (M = 10.82, SD = 3.89, SEM = 0.34, range 0-18, ICC = .99); Cooke Factor 1 (M = 4.89, SD = 2.09, SEM = 0.18, range 0-8, ICC = .70); Cooke Factor 2 (M = 4.41, SD = 2.37, SEM = 0.21, range 0-8, ICC = .88); and Cooke Factor 3 (M = 6.73, SD = 2.42, SEM = .21, range 0-10, ICC = .78).

Correlations between indices of the HCR-20 and PCL-R ranged between .20 (p < .05; Historical scale and Cooke Factor 1) and .81 (p < .001, HCR-20 Total and Hare Factor 2). All correlations were significant at least at the p < .05 level.

The relation between mean HCR-20 and PCL-R total scores and conviction for different types of past violent and non-violent crimes was examined. Both measures demonstrated a similar pattern on these crime characteristics. Compared to inmates who had not been convicted of past murder, women with such convictions scored significantly lower on the PCL-R (M = 19.77 vs. 24.22), p < .001, and HCR-20 (M = 17.50 vs. 21.69), p < .01. In contrast, compared to inmates who had not been convicted of past property crimes, women with such convictions scored higher on the PCL-R (M = 25.66 vs. 21.76) and HCR-20 (22.95 vs. 19.17), both ps < .01.
Minor crimes showed a similar pattern to property crimes, with women who had such past convictions scoring significantly higher than women without such convictions on the PCL-R (24.23 vs. 19.17) and HCR-20 (22.05 vs. 16.16), both ps < .001. The only other significant difference observed was that women who had been convicted of a past potentially violent crime had higher PCL-R scores than women without such convictions (24.75 vs. 22.03, p < .05). The HCR-20 did not differentiate women with and without past convictions for potentially violent crimes. Neither measure significantly differentiated women in the remaining community crime categories (i.e., violent, sex, and drug). Further, there were no significant differences between high and low scorers on either measure in terms of whether they had been involved in previous institutional (prison) violence. In terms of the HCR-20 scales, the only crime category in which a significant difference was observed was for minor crimes: women with such convictions had higher mean scores on the Historical scale (12.11 vs. 9.00, p < .05).

ROC analyses demonstrated a similar pattern of results for the HCR-20 and PCL-R in predicting various types of past criminal charges. Both measures were most accurate in predicting minor charges: HCR-20 (AUC = .74, SE = .05, 95% CI: .64-.84, p < .01), PCL-R (AUC = .71, SE = .05, 95% CI = .61-.81, p < .01). For both measures, the lowest AUC value obtained was for past first-degree murder: HCR-20 (AUC = .30, SE = .05, 95% CI: .20-.41, p < .01); PCL-R (AUC = .30, SE = .05, 95% CI: .20-.41, p < .01). That is, higher scores on both measures were a better predictor of not having first-degree murder charges. Finally, results (which may seem somewhat counterintuitive) demonstrated that neither measure was significantly better than chance in predicting violent charges: HCR-20 (AUC = .49, SE = .05, 95% CI: 0.38-0.59, p = ns), PCL-R (AUC = .46, SE = .05, 95% CI = .36-.56, p = ns). The authors discussed the implications of these findings for rehabilitation and treatment.

We note that all analyses involved analyses of the relationship between the HCR-20/PCL-R and past crime and violence, rather than future crime and violence.

Abridged abstract (English translation of the study not available):

Violence management and prevention exceed the exclusive court intervention and require other professionals, such as psychologists, who can help in specialized tasks like dangerousness assessment and violence recidivism control. The latest improvements in the prevention of violence have proposed the replacement of dangerousness assessment for the violence risk assessment. This new technology is more efficient to predict the future violent behavior. In this study, we present the basis of these techniques for the violence risk assessment, as well as the Spanish adapted tools and instruments for its application in clinical, forensic and correctional psychology. These are, briefly described, the VRAG scale, the PCL-R, the HCR-20, the SCR-20, the SARA test, the EPV and the SAVRY, as well as their specific functions and applications.


Summary

This study examined the rate and nature of violent acts committed in correctional and forensic settings by individuals with a diagnosis of severe mental illness, as well as the clinical and descriptive characteristics of participants who were violent during their stay. The study sample consisted of 178 (137 men and 41 women) from forensic and correctional settings. For each participant, the SCID-I, SCID-II, PCL-R, HCR-20, PANSS, and MacArthur questionnaire was completed. Information regarding substance use and police and institutional records were also obtained. The number and nature of institutional violence acts were recorded using the Overt Aggression Scale. Of the sample 56.7% was involved in institutional aggression, however only a small proportion of these participants engaged in interpersonal aggression.


Summary

The present study used multivariate profile analysis to identify clusters of patients with severe mental illness (SMI) who displayed similar patterns of violence and similar correlates. Participants were recruited from provincial and federal prisons and forensic and civil psychiatric hospitals in Eastern Canada. The final sample consisted of 178 (137 male and 41 female) participants who had been diagnosed with one of following disorders: schizophrenia, schizoaffective disorder, schizophreniform disorder, delusional disorder, a psychotic disorder not otherwise specified (NOS), major depression, or bipolar disorder. In addition to a diagnosis of SMI, participants also had a history of interpersonal violence.
Data for the study were collected from participant, criminal, and institutional records by trained professional or graduate student research assistants. All participants completed an interview that included questions about sociodemographic variables and the SCID I and II, PANS, PCL-R, HCR-20 and MacArthur Community Violence Instrument. Twenty-five participants were assessed independently by two raters in order to estimate interrater reliability of diagnoses, and risk instruments. Agreement was high for both primary diagnoses (Kappas ranged between .65 and 1.0) and comorbid disorders (Kappas ranged between .73 and .92). Intraclass coefficients were also high for the PCL-R (.94) and the HCR-20 (.92). Alpha (α) coefficients were .97 for the PCL-R total score and .83 for the HCR-20 total score, respectively.

Types, locations, methods and victims of violence, psychopathy traits, and impulsivity were entered into multiple correspondence analyses (MCA). The MCA generated a solution identifying four dimensions explaining 64% of total inertia. From these four dimensions, a hierarchical cluster analysis extracted four distinct subgroups of participants. Cluster I (“Psychotic offenders”) classified 18% of the participants, who were more likely to have been recruited in forensic wards, reported a homicide and to have used a weapon, against a family member, in their residence, when they were experiencing delusions and hallucinations and not substance intoxication.

Cluster II (“Repetitive violent offenders”) also included 18% of the participants. All members of this cluster had acquired three or more convictions or judgments of non-responsibility for nonviolent crimes and one or more conviction(s) or judgment(s) of non-responsibility for violent crimes. Almost half the participants in this cluster reported three or more lifetime violent incidents and one-third had records of two or more incidents of violence within institutions. More of the participants in this cluster had used weapons aimed at a friend or a stranger, in a public place. The violent behavior of participants in this cluster was more often associated with drug and alcohol intoxication and less often with delusions or auditory hallucinations. Compared with the other clusters, members of Cluster II obtained higher PCL-R, HCR-20, and impulsivity scores and were more likely to have a diagnosis of alcohol and drug abuse/dependence, and APSD. Cluster II included greater proportion of participants who had been recruited within correctional facilities.

Cluster III (“Institutional violent behavior”) included 24% of the sample. These participants reported few incidents of lifetime violence but a significantly higher number of incidents of violence in the previous five months compared with the other subgroups. Violent behavior in this cluster was generally not related to alcohol or drug intoxication, it never involved a weapon, it was more likely to be directed toward acquaintances (e.g., health care providers) or strangers than family members, and it occurred most often in an institution. Members of Cluster III were found equally in correctional facilities, forensic wards and involuntary hospital commitments. Cluster III obtained the highest PANS positive scale score among all participants around the time of release

The last cluster (“Less violent and stabilized”) included the largest proportion of participants (40%). Compared with the other three clusters, participants in Cluster IV had the lowest scores on the HCR-20 and reported the lowest number of lifetime and recent incidents of violence, and, consequently, they also reported the lowest number of incidents of violence when intoxicated, delusional, or experiencing hallucinations. They also had the lowest number of convictions or judgments of non-responsibility for any type of crime and for violent crimes. Overall, the violence of the participants in this cluster was less frequent, less severe and had occurred in the past. Cluster IV included proportionately more women (34%). A significant proportion of that group reported having exclusively been the victim of past violent acts (27%).

For the sample as a whole, average score on the HCR-20 was 21.9 (SD = 7.1). Average HCR-20 scores were 20.4 (SD = 5.6), 27.7 (SD = 5.2), 21.8 (SD = 7.2) and 20.0 (SD = 7.1) for Cluster 1, Cluster 2, Cluster 3 and Cluster 4, respectively. The authors concluded that among persons with SMI and considered at risk for violence, there are subgroups who display distinctive patterns and correlates of violent behavior. The finding that factors promoting violence in the subgroups differ indicates the need for distinct treatment and management strategies to reduce violence in each type.

Project and Scholarly Work


Summary

Côté observed that much research on the HCR-20 has been carried out in very different settings (i.e., civil psychiatric; forensic psychiatric; severe offenders; general offenders) and set out to evaluate the performance of the HCR-20 as a function of such settings. They sampled from involuntarily
committed patients, forensic patients who had been found not criminally responsible on account of mental disorder, and mentally disordered inmates, all of whom resided in institutions throughout the Canadian province of Quebec. Although the study was prospective, the data reported in this presentation were postdictive. Côté used the French translation of the HCR-20. Participants (n = 79 after attrition and refusal factored in, 68 with criminal record data) were evaluated just prior to release into the community. Although both males and females were sampled, the small number of females prompted Côté to drop them from analyses for this presentation.

Interrater reliability, based on a subset of 20 cases, was reported for the H and C scales, respectively, as follows: ICC \(_1\) = .88, ICC \(_2\) = .93; and ICC \(_1\) = .71, ICC \(_2\) = .83. Alpha was reported to be .93 and .83, respectively, for H and C.

The H scale (with H1 removed) varied significantly across groups with no previous offences (\(M = 8.43; SD = 3.10\)), non-violent offences (\(M = 9.17; SD = 3.64\)) and violent offences (\(M = 13.06; SD = 3.36\)). The Cohen’s \(d\) between the violent group and non-violent offence group was large (\(d = 1.11\)), as it was for the difference between the violent group and the no offence group (\(d = 1.43\)). The difference between C scores across these groups was not significant, although the effect was large for the violent versus no offences comparison (\(d = .73\)) (it was small – .29 – for the violent offences versus non-violent offences comparison).

The H scale, without H1, also was larger among correctional inmates (\(M = 14.32, SD = 2.87\)) than forensic patients (\(M = 11.42, SD = 3.50\)), or involuntary patients (\(M = 10.32, SD = 3.76\)) – which is consistent with other research when compared across studies. Côté reported multiple comparison correlation coefficients (eta) of .54 for the H scale, .24 for the C scale, and .50 for the H scale without H1 with respect to offence group, and .4 (H scale) and .16 (C scale) within the legal status groups.

Côté reported AUC values, using Statistics Canada’s definition of violence, of .83 (H), .81 (H without H1), and .61 (C), and .77 (H), .76 (H without H1) and .49 (C) for a more “stringent” definition of violence.

The author claimed that the findings offered support for the HCR-20 in terms of its interrater reliability and validity of the H scale across diverse groups. However, the C scale did not differentiate between groups. The author commented that this is not surprising given that it is meant to measure current dynamic factors, and the outcome measures in this study were all in the past.

### Project and Scholarly Work

#### Project and Scholarly Work


#### Summary

This study was a re-analysis of data that exist from three samples (two cited in this bibliography) for the specific purpose of assessing the degree of change in the Clinical and Risk Management scale and item scores across time and repeated assessments/codings. There were two forensic samples (Belfrage, unpublished raw data; Douglas et al., 1998) and one civil psychiatric sample (Douglas et al., 1999).
In Sample 1 (n = 193 civil psychiatric patients), it was possible to compare C scale and item scores at admission and discharge. Each item declined significantly, and the total score declined from 7.21 to 4.05. All drops in scores were large, as assessed by Cohen’s d (ds ranged from .89 to 1.75). At admission, 48% of the sample scored in the 8 to 10 range; at discharge, only 3% did so. In Sample 2 (n = 175 forensic patients), all C and R items declined, although the drops were not as large as in Sample 1. For the C Scale total score, Cohen’s d (.36) indicated a smallish size drop, and for the R Scale, a moderate sized drop (.50). In Sample 3, the C scale did not decline, but the R scale did so moderately (d = .44). It is possible that in Sample 3, being drawn from a Swedish forensic facility, that the patients were not as acutely disturbed upon admission (in Sweden there is no such concept as “Not Guilty by Reason of Insanity” and people are “sentenced” to treatment in the hospital somewhat liberally), and hence change was not observed.

These findings support the conceptualization of the C and R Scales as dynamic (changeable), and hence as appropriate targets for risk management and violence reduction interventions. The fact that the scores changed without direct efforts to change specific HCR-20 factors suggests that declines may be greater with intervention strategies tailored to dynamic HCR-20 risk factors.

See also


Project and Scholarly Works


Summary

This study used a prospective design to examine factors that predicted community violence in discharged mental patients. They also compared the contribution of relatively stable risk factors measured at a baseline period with more dynamic factors measured at different time points. The study used 129 discharged patients (75 of whom were male and 37 were female) and collected data at eight and 24 weeks after discharge. 34 of the patients were forensic cases and 78 were non-forensic. For this study, the authors created three violence categories: physical violence (level 1), any other violence (level 2) and any violence (any level 1 or 2 violence). Level 1 types of violence consisted of: hitting with fist, beating someone up, physically forcing sex on someone, threatening with weapon in hand, using a knife or firing a gun, any violence which results in injury. The authors used three types of factors to assess risk. They were: static factors (PCL-SV, H-scale of HCR-20, VRAG, VRS, MAST, age and DAST), dynamic trait factors (PBQ, NAS, BIS, interpersonal CIRCLE subscales), and dynamic state factors (BPRSS, GAF, BVC, HRS, WARS, and Psychotic rating scales for TCO symptoms and hallucinations). Significant AUCs for the static factors regarding any violence were: PCL-SV interpersonal (.64; p < .05), PCL-SV social deviance (.66; p < .01), PCL-SV total (.67; p < .01), VRAG total (.63; p < .05), HCR-20 H-scale (.62; p < .05). Significant AUCs for the static factors regarding level 1 violence were: PCL-SV interpersonal (.68; p < .05), PCL-SV social deviance (.67; p < .05), PCL-SV total (.69; p < .01), VRAG total (.66; p < .05), and HCR-20 H-scale (.68; p < .05). Significant AUCs for the dynamic state factors regarding any violence were: BPRS total (.61; p < .05), WARS (.62; p < .05). Significant AUCs for the dynamic state factors regarding level 1 violence were: BPRS total (.67; p < .05) and BPRS hostility-suspicion (.72; p < .01).

The scores on the HCR-20 changed significantly across time periods (p < .001). The predictive validity for the HCR-20 regarding any violence was: baseline (.65; p < .01), discharge (.80; p < .001) and eight week follow-up (.69; p < .01). The predictive validity for the HCR-20 regarding level 1 violence was: baseline (.63; ns), discharge (.80; p < .001) and eight week follow-up (.73; p < .01).

Regression equations for predicting any violence showed that without the HCR-20 C and R scales included in the model, 72.3% of patients were correctly classified, but with the C and R Scales included in the model, the number correctly classified increased to 85.7%. However, regression equations for predicting level 1 violence showed that without the C and R scales (85.7% correct prediction) was no different than having them included in the model (84.8% correct prediction).


Summary
In this prospective study of community violence, 129 inpatients in England were assessed prior to discharge from forensic and non-forensic psychiatric services using the HCR-20, PCL:SV, VRAG, Novaco Anger Scale (NAS), and BIS. The majority of patients were White (n = 104, 93%) men (n = 75; 67%). The mean age of the sample was 40 years (SD = 11.5). Most patients (n = 78; 70%) were diagnosed either with schizophrenia-spectrum disorder or bipolar disorder.

Nursing staff familiar with the participants were interviewed to gather collateral information to score the risk measures. Because roughly half of the sample did not have a criminal history, VRAG ratings technically were not completed using standard procedures, although the authors of the VRAG have themselves published studies using modified versions of the VRAG (Harris, Rice, & Camilleri, 2004), and hence this approach seems justifiable. The intraclass correlation coefficient (ICC) for the HCR-20 Historical scale between two researchers based on 20 cases was 0.97. For the Clinical and Risk Management scales, ICSSs between three raters based on seven cases were 0.85 and 0.83, respectively.

Patients (n = 112) were followed-up in the community on average 24 weeks after discharge to assess whether violence had occurred. Violent behaviour included any acts that resulted in physical injury, sexual assaults, aggressive acts that involved the use of a weapon, and threats made with a weapon in hand (i.e., using the MacArthur definition of violence). Base rates of violence varied as a function of measurement method. Using only official records, 10 participants (9%) were identified as having committed a violent act. In terms of self-report, 12 participants reported 16 acts of violence and 15 (13%) collateral informants reported 46 acts of violence. When the data sources were merged, the base rate of violence increased significantly to 19%.

The base rate of violence did not differ significantly between the groups of forensic and non-forensic patients. There were no significant differences between violent and non-violent groups in terms of age, gender, ethnicity or personality disorder diagnosis. Violent patients (n = 21) had significantly higher scores than the nonviolent patients (n = 91) on all measures. Mean scores for the violent vs. nonviolent patients, respectively, were as follows: HCR-20 Historical Scale (12.71 (SD = 3.87) vs. 10.30 (SD = 4.36); d = .58); PCL:SV total (13.43 (SD = 3.87) vs. 9.77 (SD = 5.65); d = .76); VRAG total (2.29 (SD = 10.01) vs. -3.64 (SD = 12.40); d = .53). The largest area under the curve (AUC) was for the HCR-20 total score (0.80). AUCs for other measures were: HCR-20 Historical Scale (0.68); PCL:SV total (0.69); VRAG (0.66); BIS (0.72); NAS (0.71).

Logistic regression analyses were completed to examine the relative contribution of the HCR-20 Clinical and Risk Management scales. The variables entered on the first model were those that demonstrated the most significant differences in the univariate and receiver operating characteristic analyses (i.e., PCL:SV total, HCR-20 Historical scale (which was entered with the PCL:SV item removed), VRAG total (without the PCL:SV item), BIS cognitive sub-scale and NAS cognitive sub-scale). In this model, only the BIS and NAS cognitive sub-scales independently predicted violence with significant odds ratios (1.18 and 1.11, respectively). When the scores on the HCR-20 Clinical and Risk Management scales were added, only these scales independently predicted community violence post-discharge, indicating that the HCR-20 Clinical and Risk Management scales added significant incremental validity to the baseline measures (although the proportion correctly classified only increased from 86% to 88%). Moreover, even when logistic regression analyses were conducted to control for several possible confounding variables (i.e., age, gender, length of inpatient stay and forensic status), the HCR-20 total score significantly predicted post-discharge violence.

Findings suggest that although risk measures based on historical factors are important for assessing violence risk, patients’ current functioning on dynamic factors that relate to mental illness and risk management significantly improve predictive accuracy.

### Project and Scholarly Work


### Summary

This study provides descriptive data on a risk assessment pilot program that was implemented in September 2001 concurrently with the establishment of the Court of Penal Execution (a special jurisdiction court in Buenos Aires). The sample is a subset of the 1,370 cases admitted to the Court during the first 18 months (September 2001 to February 2002). Of these cases, 105 were serving a sentence either in prison or a forensic psychiatric unit. The sample comprises all cases who were candidates for conditional release (N = 65). Participants were men with a mean age of 27.7 years (SD = 7.8). There were 55 (85%) convicted offenders and 10 (15%) insanity acquittees. The majority (61.5%) reported a history of severe drug abuse. A major mental illness was diagnosed in 6% of the sample. Of the
participants who were convicted offenders, the evaluation occurred, on average, 1.6 years before their prison term was completed.

The HCR-20, PCL-R, and VRAG were completed using an extensive information gathering process that included criminal records and court files, diagnostic interviews with respondents, interviews with family members, and participant interviews with a clinical practitioner when deemed necessary. The person(s) who completed the assessments was not reported. The mean HCR-20 total score was 18.58 (SD = 7.63; range 2-34). The mean PCL-R total score was 20.57 (SD = 9.05; range 1-37). The mean VRAG score was 12.17 (SD = 10.87; range -13-37).

No significant correlations between any risk measure and demographic variables were found.

**Project and Scholarly Work**


Abridged abstract (English translation of the study not available):

With its verdict in May 2011 the German Federal Constitutional Court declared the current law for preventive detention unconstitutional and obliged the legislative bodies to undertake a freedom- and treatment-oriented reform. Psychiatrists and psychotherapists are bound to provide therapeutic concepts. Currently there is a lack of information on the intended clientele. In our study we examined 26 persons serving preventive detention, 32 regular prisoners and 29 non-delinquent probands. The groups were matched according to age and intelligence. We gathered sociodemographic data, criminal records and conducted the tests SCID I, SCID II und PCL-R, K-FAF and BIS-11 to obtain diagnoses and characteristics. Based on this information, the HCR-20 and GAF were performed. In comparison to regular prisoners and non-delinquents, the group of those serving preventive detention is characterised by medium to advanced age, antisociality, psychopathy, substance abuse or addiction, aggressivity, a strong criminal record, years of imprisonment, insufficient educational and vocational training and a high risk of recidivism. In our examination of persons serving preventive detention, we demonstrate that this clientele is a group of recidivists difficult to treat. The current laws and a lack of early intervention programs have prevented and delayed their timely and possibly successful treatment. From a psychiatric point of view, there is a strong need for new therapeutic concepts to meet this challenge.

**Summary**

This study examined the predictive accuracy of the PCL-R, HCR-20, SVR-20 and the Static 99 in a subgroup of sexual offenders. These instruments were coded retrospectively based on information from forensic psychiatric court reports in sample of 90 released male sexual homicide offenders (out of an original sample of 166) in Germany. Follow-up information about criminal reconvictions (homicide, sexual offense, nonsexual violence offense and other nonsexual and nonviolence offense) was obtained from federal criminal records. The recidivism rates were 15.6% for sexual, 28.9% for violent (including sexual) and 47.8% for general recidivism in the sample. The average age of the sample at time of release was 38.78 years (SD = 11.37). Offenders had spent an average of 12.2 years in custody (SD = 7.6). While 75.6% of the released offenders had served a prison sentence, 18.9% had been detained in forensic psychiatric hospitals.

Data for scoring the risk instruments was available for all 90 offenders for the release follow-up sample as well as 49 offenders who had not yet been released from custody. For both subsamples together (n = 139) the Static-99 mean raw total score was 5.54 (SD = 1.96), the SVR-20 was 24.02 (SD = 6.52), for the HCR-20 was 18.27 (SD = 4.45), and for the PCL-R was 17.21 (SD = 8.35). For the 90 offenders of the released subgroup the Static-99 mean raw total score was 5.20 (SD = 1.86), for the SVR-20 was 22.40 (SD = 6.09), for the HCR-20 was 15.70 (SD = 4.45), and for the PCL-R was 15.72 (SD = 8.29). For the 49 offenders who had not yet been released the Static-99 mean raw total score was 6.16 (SD = 1.99), for the SVR-20 was 27.00 (SD = 6.27), for the HCR-20 was 19.80 (SD = 5.89), and for the PCL-R was 19.94 (SD = 7.83).

 Interrater reliability of the total scores of the risk assessment instruments, as indicated by values for ICC, ranged between...
Predictive validity, as indicated by AUC values, was provided for total and subscale scores of each of the four measures and each recidivism outcome. For the HCR-20 AUCs were reported for the H, C, and R scales, and the HC composite. For three outcomes, sexual recidivism, sexual or nonsexual violence recidivism and general recidivism, neither one particular risk assessment instrument nor any of the subscales showed a significant predictive result. Only for recidivism with a nonsexual violence offenses did aggregated total scores on the HCR-20 (without R-items) and the SVR-20 respectively show moderate validity increases (AUCs .66). Amongst the subscales, the HCR-20 clinical subscale also obtained a moderate level of predictive accuracy for violent, non-sexual recidivism (AUC = .69). AUC values for the aggregated HCR-20 total score were .51, .53 and .55 for general recidivism, sexual recidivism and recidivism with a sexual and or violent offense, respectively.

Analyses of the bivariate correlations between individual items in the four instruments and recidivism categories were also conducted by the authors. With respect to the HCR-20, they found that item C2 (negative attitudes) correlated positively with nonsexual violent recidivism ($r = 0.28, p < .01$). Further item H6 (major mental illness) correlated negatively with general recidivism ($r = -0.23, p < .05$).

The authors note several limitations to the study which may have resulted in low predictive validity of the measures. These include: missing information about influences during release, small sample sizes, as well as the possibility that risk assessment instruments investigated may be valid for general sex offender samples, but not for the particular subgroup of offenders with sexually motivated homicides examined in the study.


Summary

This retrospective study investigated characteristics that differentiated sexually motivated murderers that targeted child victims (CV; $n = 35$) from those with only adult victims (AV; $n = 100$). Differences were examined across several domains, including psychiatric disorders, childhood development, criminal history and risk of criminal recidivism in relation to the offenders’ specific type of victim. Three trained raters coded the SCID-II, PCL-R, HCR-20, SVR-20 and Static – 99 blind to recidivism outcomes using information obtained from the participants’ psychiatric court reports. Data on duration of detention and reconviction rates were obtained from German federal criminal records. Reoffending was limited to three types: sexual offenses, nonsexual violent offenses, and nonviolent offenses.

The study sample consisted of 166 men who committed a sexual homicide between 1945 and 1991. A majority (81.32%) of the sample was white, German (98%), single (71.1%) or divorced/living apart (15.6%) at the time of their index offense. The murderers’ mean age at the time of the first sexual homicide was 26.9 ($SD = 8.4$). Average IQ of the participants was 101.5 ($SD = 13.8$). The CV group more often met the diagnostic criteria of pedophilia (43% vs. 4%) and less often alcohol abuse and drug dependency (31% vs. 55%). Sexual dysfunctions (9% vs. 29%) and narcissistic personality disorder (0% vs. 13%). Childhood development and trauma did not significantly differ between groups. While the CV group was detained more frequently in forensic psychiatric hospitals (59% vs. 26%), the two groups showed the same rates of release and reconviction for sexual (22% for both groups), nonsexual violent (CV 25% vs. AV 15%) and nonviolent offenses (CV 63% vs. AV 59%). The high prevalence of violence (CV 31.4% vs. AV 55% and antisocial personality disorder in both groups (CV 25.7% vs. AV 24.0%) may have accounted for the similarities in rates of sexual homicide between groups.

No significant differences between the two groups were found on the risk assessment instruments. With regards to the CV group, mean scores were 15.74 ($SD = 6.89$), 16.91 ($SD = 4.32$), 22.74 ($SD = 5.98$) and 5.7 ($SD = 2.00$) on the PCL-R, HCR-20, SVR-20 and Static-99, respectively. With regards, to the AV group, mean scores were 16.39 ($SD = 6.68$), 16.15 ($SD = 4.82$), 22.83 ($SD = 6.64$) and 5.19 ($SD = 1.80$) on the PCL-R, HCR-20, SVR-20 and Static-99, respectively.

One limitation of this study pertains to the quality, quantity, and nature of information available from the 1940s compared to more recent times.

Project and Scholarly Work

**Summary**

The current study was carried out to determine the validity of risk assessment tools in Scotland. Specifically, the authors were interested in determining whether the PCL-SV, VRAG and HCR-20 were useful risk assessment tools for use with Scottish populations. Participants were 96 patients discharged from a medium secure psychiatric unit that houses patients from correctional institutions, other psychiatric units and the community. Most of the sample was male (91.7%), with a mean age of 34.2 years ($SD = 9.98$). This study used a retrospective file review design. As such, only the H scale was scored. Violence was assessed for 24 months. Of the participants, 40.6% perpetrated any violence in the follow-up period with 4.2% perpetrating a serious violent incident. Mean scores on all the assessment tools were as follows: PCL-SV = 11.48, VRAG = 5.46, and H-10 = 13.14.

For the prediction of minor violence incidents, serious violent incidents and any violence, respectively, the AUCs were .638, .664, and .625 using the PCL-SV, .697, .738, and .681 using the VRAG, and .619, .739, and .605 using the H scale. For the prediction of minor violence charges or convictions, serious violent charges or convictions and non-violent charges or convictions, respectively, the AUCs were .627, .385, and .519 using the PCL-SV, .640, .594, and .591 using the VRAG, and .605, .538, and .512 using the H scale.

The authors concluded that the mean scores and AUCs found were comparable to other studies. The VRAG appeared to be the best predictor in the present study; however, the design features may have given this tool an advantage. The utility of the three instruments in Scottish settings was further discussed.

**Project and Scholarly Work**


**Summary**

This was the first research to study the HCR-20 in Hong Kong and in a predominately Chinese population. Using a prospective cohort design, the HCR-20 was completed for 110 discharged general adult and forensic psychiatric patients with a history of criminal violence or a disposition towards violence (PFU group) and 110 demographically matched controls. Post-discharge violence was determined from case notes at 6- and 12-months after HCR-20 ratings were completed. Acts of violence were divided into three categories including physical violence, sexual violence and non-physical violence.

Participants had a mean age of 43 with a majority suffering from schizophreniform disorders (63.6%). A hundred and nine of the participants in the PFU group had a history of violence, with 71 participants having a violent conviction. However, there were also 73 participants in the non-PFU group who had a history of violence, with 31 having a past violent conviction. Within 6 and 12 months after discharge, the proportions of participants who were violent against people were 6.3% and 11%, sexually violent were 2% and 2.5% and verbally violent were 13.2% and 19% respectively. Overall, the PFU group committed more post-discharge violence than the non-PFU group at both time points.

The mean total HCR-20 score for the PFU group was 19.06 ($SD = 6.99$) and the mean total score for the non-PFU group was 14.12 ($SD = 5.64$). Means for the subscales were also reported for each group. Overall, the combined mean scores for both groups were 16.59 ($SD = 6.80$), 8.37 ($SD = 3.98$), 4.25 ($SD = 2.30$) and 3.98 ($SD = 2.19$) on the total, historical, clinical and risk management scales respectively.

Inter-rater reliability was calculated for both total scores and individual items on a total of 110 cases. ICC values were .71, .43, .37, .57 and .73 for historical, clinical, risk, total scores and final risk judgments, respectively. Predictive validity of the HCR-20 scales was established with Receiver Operating Characteristics (ROC). AUC values for the HCR-20 structured final risk judgment were significant for all outcomes including verbal violence (AUC = .67), violence against others (AUC = .78), violent conviction (AUC = .68) and any violence (AUC = .70) within 6 months. For verbal violence, violence against others and violent conviction as well as any violence, AUCs for historical and clinical subscales as well as HCR total score were all statically significant (AUCs range .67 to .80). When considering the different outcomes within 12 months, structured final risk judgments were also significant for all outcomes (AUCs = .64, .70, .67 and .67 for verbal violence, violence against others, violence conviction and any violence, respectively) and the AUC for the historical and clinical subscales and the HCR total score were significant for violent conviction (AUCs range .72 to .80) The AUC for the risk management scale was not significant for any of the outcomes, whereas
the AUC for the other subscales and HCR total score were significant for all outcomes (AUCs range .62 to .68). PFU status was able to predict verbal violence, violence against others, violent conviction and any violence greater than chance probability at both time points.

Cox regression analyses were conducted using any violence within the 12 months as the outcome. Scores on the subscales were entered on the first block and the final risk judgment was entered on the second block using the forward conditional method. The results indicated that Clinical and Risk Management scores, as well as the structured final risk judgment were significant predictors in the model. Further, the addition of structured final risk judgment significantly improved the model’s fit, ($\chi^2 (1) = 4.21) = 22.9, p < .05$).

This was one of the first studies to validate the use of the HCR-20 in a group of general adult and forensic Chinese psychiatric patients discharged into the community. Findings supported that structured final risk judgments added incremental validity to the numerical use of the HCR-20.

See Also


Project and Scholarly Works


Summary

This study presents a descriptive overview of a multi-site, international, prospective study concerned primarily with the community aftercare treatment of forensic and civil psychiatric patients. One of the stated goals of the project is to attempt “to validate the HCR-20 in four different cultures,” namely, Canadian, Swedish, German, and Norwegian. Included in this goal is the evaluation of whether there are subtypes of patients for whom the HCR-20 is less accurate, and hence might need revision. The study involves eight data collection sites, two each (one civil psychiatric, one forensic psychiatric) from the four countries.

Reliability analyses to date have included the ratings of four patients each by seven clinicians from different countries. The ICC1 values were as follows: HCR-20 Total Score (.90); H Scale Score (.94); C Scale Score (.89); R Scale Score (.68). The authors stated that “ICCs are generally very high, indicating excellent inter-rater reliability.”

At present, the project is not far enough along to provide predictive validity analyses. However, criterion-related validity has been partially evaluated through correlating the seven clinicians’ ratings with those made by the authors of the HCR-20 on the four cases. The ICC1 values for these analyses were as follows: HCR-20 Total Scale Score (.99); H Scale Score (.85); C Scale Score (.99); R Scale Score (.96).

Future research reports from this study will provide information on predictive validity in both forensic and civil psychiatric patients, from four countries, on ratings made prospectively through both clinical interview and file review methodology.


Summary

These authors used a prospective design to determine if any markers at time of discharge (such as PCL-R or HCR-20 scores) would be predictive of future violent acts. They used a male only sample with schizophrenia spectrum disorder and followed 128 individuals up to 6 months after discharge from either forensic (aggressive individuals) or general psychiatry (non-aggressive individuals) clinics and
then they were able to follow up 78 of the original 128 for a second six month period to make it a 1 year follow up for this smaller group. This report was based on the Hodgins et al. (2007) sample reported above.

At discharge, the aggressive individuals showed a mean of 15.1 (SD = 8.1) on the PCL-R and a mean of 22.2 (SD = 6.5) on the HCR-20 score. The HCR-20 subscale means were: H-scale mean 12.8 (SD = 3.7), C-scale mean 3.9 (SD = 2.8) and R-scale mean 5.6 (SD = 2.2). At discharge, the non-aggressive individuals showed a mean of 12.2 (SD = 7.2) on the PCL-R and a mean of 17.2 (SD = 6.5) on the HCR-20 score. The HCR-20 subscale means were: H-scale mean 10.4 (SD = 4.4), C-scale mean 3.1 (SD = 1.9) and R-scale mean 3.8 (SD = 2.2).

Results showed that the PCL-R did not predict aggressive behaviour. Results also showed that the HCR-20 H-scale did not predict future violent acts. However, the C and R scales did predict future aggressive behaviour. The strongest predictors of future violent behaviour in this sample were increases in anxiety and depressive symptoms over time.

Logistic regression analyses to predict aggressive behaviour during the first follow-up period indicated the HCR-20 total score increased risk by 1.1 times per unit and the R-scale increased risk by 1.5 times per unit (TCO and a score of 5+ on PANSS positive symptoms increased risk by 1.2 and 5.2 times, respectively). TCO and HCR-20 indices did not remain significant when PCL-R total score and diagnosis of substance abuse/dependence were controlled for. For the second follow-up period, odds ratios for HCR-20 total, C-scale, and R-scale were 1.2, 2.1, and 2.2, respectively. These values are substantially lower in comparison to values for the other measures (e.g., a score of 5+ on PANSS positive symptoms yielded an odds ratio of 34.0). As was the case for the first follow-up period, no HCR-20 indices remained significant once PCL-R total score and diagnosis of substance abuse/dependence were controlled for.


Summary

This study examined the predictive validity of the HCR-20 total, subscale, and final risk ratings for aggressive behavior among patients with schizophrenia under the care of forensic and general psychiatric service. Patients were followed up with for a 2-year period. Physical aggression was determined using the MacArthur Community Violence Interview which collected information from the participant as well as a collateral source. Data was collected prospectively and patients were assessed at baseline and at 6, 12, 18 and 24 months thereafter. Patients could miss one follow-up but complete the subsequent follow-up.

The study sample was comprised of 248 male patients (150 from forensic hospitals and 98 from general hospitals) hospitalized in Canada, Finland, Germany, and Sweden. Of the sample, 84% were diagnosed with a schizophrenic disorder. Rates of APSD and prior nonviolent/violent convictions were higher in the forensic sample compared to the general psychiatric sample. Of the sample, 24% of participants from the forensic group dropped out of the study and 31.6% from the general psychiatric group.

Inter-reliability was determined for a random sample of 35 cases. ICC values were 0.90, 0.78, and 0.52 for the H, C, and R items of the HCR-20, respectively. Scores on each of the subscales, but not the total score, significantly differed between forensic and general psychiatric samples. With regards to the forensic group, mean scores were 11.65 (SD = 3.91), 3.18 (SD = 2.11), 3.87 (SD = 2.35) and 18.71 (SD = 6.83) on the H, C, R and total scales of the HCR-20. With regards to the general psychiatric group, mean scores were 9.19 (SD = 4.50), 4.70 (SD = 1.88), 5.30 (SD = 2.07) and 19.19 (SD = 7.04). Overall, the proportion of forensic and general psychiatric patients who engaged in aggressive behavior was low; however more of the general than the forensic patients engaged in aggressive behavior over the two-year follow-up (24.5% vs. 11.3%).

ROC analyses indicated that AUC values differed between forensic and general psychiatric patients and that accuracy of prediction varied by the length of follow-up, particularly among the general psychiatric patients. Among the forensic patients, the AUCs for 24 months following discharge indicate that the total HCR-20 score, the scores for each subscale and the final risk ratings accurately predicted aggressive behavior with medium to large effect sizes (AUCs range = .66 to .74). By contrast, among the general psychiatric patients, total score (AUC = .72) and the H scale (AUC = .75) significantly predicted aggressive behavior while the predictions for the C (AUC = .59) and R (AUC = .61) score were not better than chance. Among the forensic patients, the AUCs for the 6, 12 and 18 months following discharge indicate that the H score was most accurate in predicting aggression behavior (AUCS range .62 to .76), however the AUC value for 18 months was not significant. Among the general psychiatric patients, the C and R scores did not predict aggressive behavior any better than chance over the periods of 6, 12, and 18 months (AUCS range = .54 to .66), while among the forensic patients both the C (AUC = .72) and R (AUC = .69) scores were significant predictors.
at 18 months. Among the forensic patients, the total HCR-20 score (AUCs range 67 to .74) and SPJ (AUCs range .65 to .69) provided significant predictions of aggressive behavior regardless of follow-up period. Among the general psychiatric patients however, the prediction over the first six months after discharge was not better than chance for the total score (AUC = .60), however total score did significantly predict aggressive behavior at 12 months (AUC = .74), and 18 months (.72). With respect to SPJ, AUC values were not better than chance for either the 6 months (AUC = .57) or 12 months (AUC = .63) after discharge.

The authors also assessed whether the accuracy of prediction varied depending on the length of the follow-up period and if changes in the C and R items at 6-month intervals would predict aggressive behavior in the subsequent 6 months. To investigate whether changes in C and R scores were associated in changes in the risk of aggressive behavior, the authors examined change in individual C and R items using an individual bivariate GEE analysis. These analyses determined that changes in three of the clinical risk items (negative attitudes, impulsivity and unresponsive to treatment) significantly predicted changes in aggressive behaviors (increases in risk factors predicted increases in occurrence of violence). Changes in three risk management items (plans lack feasibility, lack of personal support and noncompliance with remediation attempts) and total R scores were also significant predictors of increased violent behavior.


Summary

This study compared the rate of past criminal behaviour among male patients being discharged from forensic and general psychiatric hospitals in four countries (same study as Hodgins et al., 2007, above). Clinicians also assessed the risk of violent behaviour in the future using the Psychopathy Checklist and the HCR-20. The sample consisted of 110 forensic patients and 47 general psychiatry patients. Patients, all of whom were men, had either schizophrenia, schizoaffective, or schizophreniform disorder. Within each site, each patient being discharged from a forensic hospital was matched to a patient of the same sex, age, and primary diagnosis being discharged from a general psychiatric service.

A number of comparisons between the forensic and general psychiatric groups was undertaken (e.g., age at discharge, various criminal history indices, type and history of psychiatric admissions). In addition to criminal history variables, comparisons of a more clinical nature were made between the general psychiatric patients with and without criminal histories. Relative to general psychiatric patients without a criminal history, those with a criminal history had significantly higher HCR-20 total scores ($X = 21.00$, $SD = 5.54$ vs. $X = 15.54$, $SD = 6.18$; $t$ (47) 2.70, $p = .01$), R-scale scores ($X = 6.25$, $SD = 1.42$ vs. $X = 4.23$, $SD = 2.21$; $t$ (47) 3.64, $p = .001$), and PCL-R total scores ($X = 13.27$, $SD = 6.36$ vs. $X = 9.34$, $SD = 5.42$; $t$ (47) 2.07, $p = .04$). No significant difference between patients with versus without a criminal history was observed for H-scale scores ($X = 9.75$, $SD = 5.05$ vs. $X = 7.29$, $SD = 3.16$; $t$ (47) 1.59, $p = .13$), C-scale scores ($X = 5.00$, $SD = 1.81$ vs. $X = 4.03$, $SD = 1.98$; $t$ (47) 1.50, $p = .14$), PCL-R Factor 1 scores ($X = 4.67$, $SD = 3.42$ vs. $X = 3.43$, $SD = 2.62$; $t$ (47) 1.31, $p = .19$), or PCL-R Factor 2 scores ($X = 7.41$, $SD = 2.67$ vs. $X = 5.60$, $SD = 3.67$; $t$ (47) 1.56, $p = .13$). A global clinical judgment (it was not specified whether this judgment was based solely on information gathered during the completion of the HCR-20) of risk for future behaviour (low, moderate, high) over the subsequent 6 months did not distinguish the two groups. It was not specified whether this clinical judgment was made according to the SPJ model, or was unstructured.

The patients with a criminal history were assessed as having a greater risk for violent behaviour in the community after discharge as indicated by higher total scores on the PCLR and on the HCR-20. Lastly, the global clinical judgment of risk of future violence did not distinguish between the two study groups (forensic and general psychiatry).

Project and Scholarly Works


Abridged Abstract (English translation of the study not available):

The risk of violent behaviour after discharge from German hospitals among 50 forensic psychiatric and 29 civil psychiatric patients was studied. Participants were assessed at discharge and at four follow-ups over a period of two years. Differences in the psychopathology, the use of psychiatric aftercare, violent incidents, and the risk of acting violently according to the HCR-20 were computed. Forensic patients had more frequent contacts with
psychiatrists or clinical psychologists as well as more frequent supervised activities. Civil psychiatric patients exhibited more clinical symptoms during a six-month post-discharge period. The risk of violence decreased during the follow-up period in both groups. Forensic and civil psychiatric patient groups did not differ significantly with respect to the amount of risk attenuation observed during follow-up or to the number of violent incidents. The authors concluded that existing aftercare appears to decrease the risk of violence after discharge in both patient samples, and that the data are inconsistent with widespread fears that patients discharged from a forensic hospital pose an increased risk to others.


Abridged Abstract (English translation of the study not available):

The aim of this study was to assess the potential for future violent behaviour comparing patients recruited from forensic and general psychiatric wards in Germany. Patients and methods. Fifty patients were recruited from a forensic hospital and 29 from a general psychiatric hospital. In the weeks preceding discharge, structured assessments of the future risk of violent behaviour were completed using the HCR-20. Results. There was little difference in the risk presented by the two groups. Forensic patients presented an elevated risk of violence because of historical factors, while the risk among patients from general psychiatry was due to clinical symptoms. Conclusion. Some criminal offences could be prevented if more time and effort were spent in general psychiatric practice in identifying patients at high risk for violence and in reducing symptoms of psychoses before discharge.

---

Project and Scholarly Work


Summary

The authors investigated the VRAG, HCR-20, Static-99, and the Risk Matrix 2000 in a sample of 212 intellectually disabled offenders. Participants were distributed across three facilities with differing levels of security in England and Wales: high security (L1; N = 73), medium (L2; N = 70), low security (L3; N = 69), and a community forensic service. Participants were all available people in L1, a random sample in L2, and consecutive referrals for L3. Participants in L1 had committed more violent offences and participants in L2 had significantly less mental disorders. In addition to these assessment instruments, violent and sexual incidents were recorded over a period of 12 months, and the predictive value of each instrument is reported (pseudo-prospective study). Risk assessment instruments were coded based on clinical files and, when the information was questionable, information was gathered from relevant key workers and support workers.

Violent and sexual incidents (significant incidents) were recorded independent of the study through the nursing case notes. A significant incident was defined as verbal aggression, physical aggression, destruction of property, and inappropriate sexual behaviour. Reliability between raters was 100% for whether an incident was violent or not, sexual or not, and sexual or violent.

Inter-rater reliability was assessed on 30 cases with independent raters. For the VRAG, on Risk Categories 1–9, agreement was 92.2%. For the HCR-20, agreement was high (H subscale, 89.4%; C subscale, 93.1%; R subscale, 82.7%). For the RM 2000-V, reliability was 90.7% and for the RM 2000-S, 92.1%; for the Static-99, reliability for Risk Levels 1–4 was 97.2%.

The H subscale mean for whole sample was 12.09 (SD = 4.43). The mean for the C subscale was 4.41 (SD = 2.40), and for the R subscale the mean was 3.04 (SD = 1.78). The VRAG mean was 6.73 (SD = 8.84). Overall, the findings revealed significantly higher risk assessment scores for L1 over L3 (HCR-20 and VRAG) and L2 over L3 (HCR-20 only). Of the actuarial assessments, the VRAG, HCR-20-H, and the RM 2000-C showed a significant difference between groups. The VRAG and the H subscale significantly predicted violent incidents (AUC = .71, .72 respectively). The H, C, and R subscales all significantly predicted violent incidents (AUCs of .68, .67, .62 respectively). The RM 2000-V was not significant. For sexual incidents, the Static-99 was the only significant predictor (AUC = .71).

**Summary**

Previous authors have noted that Bayesian analysis is suitable for assessing recidivism risk because it takes into account the base rate of recidivism in terms of a prior probability. This retrospective study assessed the potential of Bayesian classification as a method for violence risk assessment within a multivariate framework using a sample of 393 German male offenders. Participants in the study were originally assessed for criminal law proceedings from the mid 1980s to mid 1990s by members of a Forensic Psychiatry Unit at a psychiatric hospital in Germany. The mean age of the men in the sample at the time of their original assessment was 33.3 years (SD = 11.93). Time at risk was calculated from three points onwards: release from prison, discharge from a high secure psychiatric hospital or the day of the trial (if the offender was sentenced to probation). The endpoint was date of a new offense leading to a reconviction or January 31, 2002 if no further convictions occurred. On average, the time at risk was 6.5 years (SD = 59.7 months). Data on reconvictions were obtained through a content-analysis of personal records from the German Federal Registry of Convictions. Outcomes were coded as one of the following: no additional convictions occurred. For a subset of eleven cases that were assessed by two raters independently ICC values were .92 and .89 for the PCL-R and HCR-20, respectively.

Analyses were conducted in two steps. First, the combination of predictor variables that worked best within a multivariate Bayesian classification (MBC) framework was identified. Predictor variables considered were: age, PCL-R Total, Factor 1, Factor 2, HCR-20 total, H, C, and R total scores. For the eight above-mentioned variables this amounted to 255 possible combinations, ranging from a single predictor to the combination of all eight. Second, the resulting optimal model for the area under the ROC curve was compared with the corresponding AUC values for the individual risk factors. Among the 255 combinations of total or subset test scores, the pair of age and Factor 2 on the PCL-R (AUC = 0.75, p < .001) was the most effective in allocating cases to the groups of recidivists or non-recidivists within an MBC routine. With respect to violence risk, AUC values for the single predictors were 0.65, 0.68, 0.62, 0.71, 0.67, 0.63, 0.61, and 0.61 (all ps < .01) for age, PCL-R total scores, Factor 1 score, Factor 2 score, HCR-20 total score, H score, C score and R score, respectively. On average, non-recidivists were older (M = 34.5 years, SD = 12.1 years) than recidivists (M = 28.0 years, SD = 9.6 years) and recidivists had a higher mean score on Factor 2 of the PCL-R (M = 12.2, SD = 4.6) than non-recidivists (M = 8.2, SD = 4.7).

The results suggest that the amalgamation of various predictors may yield significantly higher AUC values than individual predictors. The authors note that the results obtained may not be representative of other samples (such as those in North America) because of the lower rates of violent recidivism in the study sample, and differences in offender age and scores on total and subtotals of the HCR-20 and PCL-R.

**Project and Scholarly Work**


**Summary**

This study was undertaken to examine the utility of the Problem Identification Checklist – Revised (PIC-R) in predicting inpatient and community violence. The authors discuss that much research in the risk assessment field has focused on static risk factors and assessment tools. Therefore, there exists a need for the development and validation of tools design to assess dynamic risk. Using a retrospective design, participants included 49 psychiatric patients from a low security mental health facility that housed both forensic psychiatric patients and civil psychiatric patients. Of the participants, 25 were inpatients on the unit and 24 were outpatients that previously residents of the unit. With a mean age of 42 years (SD = 11.54), most of the participants were male (75.5%) and Black Caribbean (51%) or Caucasian (35%).

The H scale was scored based on file information, omitting item H7 (Psychopathy). The PIC-R was rated by nursing staff. Violent outcomes were coded from file information using the Overt Aggression Scale (OAS). Violence was separated into three categories: any physical aggression, aggression towards objects, and aggression towards people. The mean H scale score was 10.68 for the inpatient sample and 9.25 for the community sample. With regards to the inpatient sample, AUCs of .52, .53, and .57 were found for the H scale and aggression against others, against objects,
and any aggression, respectively. For the PIC-R, the AUCs were .79, .82, and .74 for each type of violence. With regards to the community sample, the AUCs for the H scale were .82, .66, and .82 for aggression against others, against objects, and any aggression, respectively. AUCs for the PCI-R were .64, .75, and .68, respectively.

The authors discussed the implications of the results, including the finding that the PIC-R showed higher AUCs in terms of inpatients violence, while the H scale showed higher AUCs in terms of community violence.

Project and Scholarly Work


Summary

This retrospective study compared HCR-20 risk profiles of forensic and civil psychiatry patients referred to forensic mental health services in Australia. Study data was gathered from comprehensive assessments completed by trained clinicians at a medium secure forensic facility and a community forensic mental health service. Data from 129 forensic patients (19 from the medium secure facility and 110 from the community service) and 105 civil patients (18 from the medium secure facility and 87 from the community service) were included. Overall, 94% of patients were male and 73.5% had a diagnosis of schizophrenia. Co-morbid substance misuse and psychotic illness was observed in 40.2% of the overall patient group. Amongst the forensic patients, homicide/attempted homicide was the most common type of index offence (57.4%).

Mean scores on HCR-20 total scores and subscales were compared between the broader civil and forensic patient groups as well as between civil and forensic patient groups within the medium secure facility (CiF & FiF) and the community (CiC & FiC). Standard deviations were not reported. The civil patient groups produced means on the HCR-20 that were consistently higher than the forensic patient groups. On the whole civil patients had a mean total score of 25.21 on the HCR-20, while forensic patients had a mean total score of 16.85. Patients in the CiC group had a mean total score of 24.82 and FiC of 16.20. Means were 27.11 and 20.58 for the CiF and FiF groups, respectively. All differences were significant. A significant difference between means was not found on the H subscale when comparing CiF and FiF groups, however significant differences between means were found on all other subscales as well as the H subscale when comparing CiC and FiC groups, and civil and forensic patients as a whole. Means produced across C and R subscales were also compared across groups. The mean from combined scores on the C and R subscales for the civilian patients was 13.99 compared to 7.20 for the forensic patients. This difference was significant. Significant differences were also observed between the CiC and FiC, and CiF and FiF groups, with civilian groups scoring higher on the combined C and R subscales.

A greater proportion of civil patients (57.1%) fell into the high-risk category (i.e. score of 25 or higher on the HCR-20) compared to 17.8% of forensic patients. 54% of patients in the CiC group fell into the high-risk category compared to 15.5% of patients in the FiC group. 72.2% of patients in the CiF group fell into the high-risk category compared to 31.6% of patients in FiF group. Odds ratios indicated that civil patients were 6.15 times more likely to fall into the high-risk category on the HCR-20 than forensic patients. The odds of CiC patients falling into the high-risk category were 6.4 compared to the FiC group. The odds of CiF patients falling into the high-risk category were 5.6 compared to the FiF group.

Project and Scholarly Work


Summary

In this prospective study, violent recidivism was compared between mentally disordered offenders sentenced to prison and those sentenced to forensic care over an average period of five years. As part of this study, the predictive ability of common criminological (e.g., age at first conviction, number of convictions for violence, number of prison convictions) and clinical risk factors (e.g., psychosis, substance abuse/dependence, APSD and conduct disorder), and structured violence risk instruments, the HCR-20, PCL-R and Life History of Aggression (LHA), were examined. The sample used in this study was comprised of 100 mentally disordered perpetrators (92 men and 8 women) of severe violent and/or sexual crimes in Sweden. Of these participants, 46 had been sentenced to compulsory forensic care and 54 to prison. Subjects were consecutively assessed between 1998 and 2001 by a clinical battery of instruments including the risk instruments listed above. Five year follow-up data was collected from official registers for recidivistic crimes (all types, particularly sexual and violent crimes) and violent recidivism (all violence-related
Reconvictions for violent criminality were significantly higher among those sentenced to forensic psychiatric care (28% vs. 11%). A Kaplan Meier survival analysis with a Log-Rank test indicated significant differences until time to violent relapse between the two groups. Those sentenced to forensic care had fewer relapses and relapses were confined to the beginning of the follow-up period, whereas those sentenced to prison displayed relapses spread out over the entire follow-up period.

ROC analyses showed modest predictive validity for age at first conviction (AUC = .70) and all three risk instruments. AUC values were 0.71, 0.74 and 0.74 for the HCR-20, PCL-R and LHA. Further, age of first conviction and each of these instruments was significantly correlated with violent recidivism during the follow up, $r = -0.28, .33, 0.29, 0.33$ for age, PCL-R, HCR-20 (total score on historical and clinical items) and LHA.

ROC analyses were also used to identify the optimal inflection point for each of the risk measures. For the HCR-20 the optimal inflection point was 12.5, where a score above that value correctly classified 63% of the studied subjects. For the PCL-R the optimal inflection point was 11.5 where a score above that value correctly classified 70% of subjects. The LHA had an optimal inflection point of 27, where a score above correctly classified 65% of subjects.

Logistic regression analyses were used to identify the best criminological or clinical/structured assessment risk predictors for violent recidivism. The results indicate that that predictive ability of criminological risk factors versus clinical factors combined with scores from assessment instruments were comparable, with each set of variables managing to correctly classify 80% of the individuals, but the only significant predictors in the model were age at first conviction and substance abuse/dependence among primary relatives. While the structured assessment instruments showed only modest relations with violence, the authors note that the findings should be interpreted with caution due to the small size of the sample which likely limited the ability to detect statistically significant associations between risk factors and outcome.

**Summary**

The overall focus of the current research was to examine the concurrent and divergent validity of the Short Term Assessment of Risk and Treatability (START). Two separate samples were used and the procedures were different for each sample. In the civil psychiatric sample ($N = 21$) patients were interviewed and their files were reviewed before the assessments were completed. The START was then coded by a researcher masked to HCR-20 and PCL:SV ratings. In the forensic psychiatric sample ($N = 68$), all measures were coded by the same researcher based on a retrospective file review. The forensic sample consisted of all male patients with a mean age of 40.6 years and the majority (85.3%) had a diagnosis of some type of schizophrenia. The civil psychiatric sample consisted of mostly males (71.4%) and again the majority (76.2%) had a diagnosis of some type of schizophrenia.

With respect to the civil psychiatric sample, mean scores on the risk assessment measures were as follows: START Strengths $M = 13.71$, $SD = 6.03$; START Vulnerabilities $M = 18.71$, $SD = 6.15$; H scale $M = 11.71$, $SD = 3.77$; C scale $M = 5.24$, $SD = 2.00$; HCR-20 Total score $M = 22.38$, $SD = 6.15$; and PCL:SV $M = 10.62$, $SD = 4.56$. With regards to the forensic psychiatric sample, mean scores on the risk assessment measures were as follows: START Strengths $M = 17.90$, $SD = 7.82$; START Vulnerabilities $M = 17.09$, $SD = 8.20$; H scale $M = 13.22$, $SD = 3.70$; C scale $M = 4.75$, $SD = 2.65$; HCR-20 Total score $M = 23.98$, $SD = 6.75$; and PCL:SV $M = 11.19$, $SD = 4.19$.

The HCR-20 total score was found to correlate with the START Vulnerabilities scale in the civil psychiatric sample ($r = .48$) and the forensic psychiatric sample ($r = .76$). The C scale alone was also found to correlate with the START Vulnerabilities scale in both the civil psychiatric ($r = .67$) and forensic psychiatric ($r = .85$) samples. At the item level, all of the C scale items correlated with their corresponding items on the START, with coefficients ranging from .64 to .78. The H scale was also found to correlate with the START Vulnerabilities scale in the forensic psychiatric sample ($r = .41$), but not in the civil psychiatric sample ($r = .17$, ns).

The HCR-20 total score was also found to correlate with the START Strengths scale in both the civil psychiatric ($r = -.54$) and forensic psychiatric ($r = -.75$) samples. The C scale

---

**Project and Scholarly Works**

also correlated with the START Strengths scale in both samples \((r = .50\) for the civil psychiatric sample, and \(r = .80\) for the forensic psychiatric sample). At the item level, again, the C scale items correlated with their corresponding items on the START, with coefficients ranging from \(-.56\) to \(-.77\). The H scale was also found to correlate with the START Strengths scale in the forensic psychiatric sample \((r = -.39)\), but not in the civil psychiatric sample \((r = -.35, ns)\).

Notably, the final risk estimates made using the HCR-20 and the final risk estimates made for violence using the START did not correlate \((r = .21, p = .36)\).

Additional correlations are reported between the START and the PCL:SV. The authors concluded that the results support the concurrent and divergent validity of the START, but future research is needed.

**Petersen, K., Nicholls, T., Brink, J., & Webster, C. (2009, June). Convergent and divergent validity of the START with the PCL:SV and HCR-20. In T. Nicholls (Chair), Implementation and evaluation of START in civil and forensic psychiatric services. Symposium conducted at the annual conference of the International Association of Forensic Mental Health Services, Edinburgh, Scotland.**

**Summary**

This presentation also reported on the convergent and divergent validity of the START with the HCR-20 and PCL:SV. Moreover, this presentation reports on overlapping samples with the study summarized above. The current study reports on the same forensic psychiatric sample as discussed above, as well as another forensic psychiatric sample \((N = 28)\). Only the results from this new sample are reported here. This sample was predominantly male \((92.9\%)\) and had a diagnosis of schizophrenia \((82.1\%)\).

Based on a retrospective file review, the mean scores on the risk assessment measures were as follows: START Strengths \(M = 9.57, SD = 5.47\); START Vulnerabilities \(M = 22.79, SD = 6.09\); H scale \(M = 11.11, SD = 4.82\); C scale \(M = 5.75, SD = 2.35\); HCR-20 Total score \(M = 24.14, SD = 8.23\); and PCL:SV \(M = 9.30, SD = 5.94\).

In this sample, the START Vulnerabilities scale was found to correlate with the C scale \((r = .46)\), but not with the H scale or HCR-20 total score. The START Strengths scale did not correlate with the HCR-20 total score, H scale, or C scale. The HCR-20 total score did not correlate with the final risk judgments made using the START.

The results of this sample and the other sample are discussed in terms of their implications for the START.

---

**Project and Scholarly Work**


**Abridged Abstract (English translation of the study not available):**

This study assessed the convergent and predictive validities of the HCR-20, PCL-R, and VRAG. Participants were drawn either from high security prisons or a forensic hospital. The three instruments correlated highly \((> 0.70)\), sharing a large common variance. Receiver Operating Characteristics, survival curves analyses and correlation coefficients suggested that the three measures presented a moderate predictive validity both for general and violent recidivism.


**Summary**

This study examined the extent to which the PCL-R and HCR-20 items varied with respect to the amount of subjective judgment used to assign scores and the extent to which perceived scoring subjectivity was associated with interrater agreement. Nine doctoral clinicians with experience scoring the two instruments rated each item with respect to the amount of subjectivity to score each item. There was a high level of agreement with respect to the amount of subjectivity required to score the items \((ICC_c = .94\) and \(.92\) for the PCL-R and HCR-20, respectively) and which had the most useful scoring guidelines \((ICC_c = .79\) and \(.87\), respectively).

To examine the association between scoring subjectivity and factor/facet-level rater agreement of the PCL-R, subjectivity composite ratings were averaged for items on the same factor or facet separately for each doctoral student clinician. Multivariate analyses of variance (MANOVAs) were used to examine whether the clinicians perceived Factor 1 and its facets as requiring more subjectivity that
Factor 2 and its facets. The first MANOVA showed that clinicians saw Factor 1 items as requiring more subjectivity (Cohen’s $d = 2.26$) than Factor 2 items. The second MANOVA showed that there was a significant difference in ratings for the four facets. Items on Facet 4 were seen as requiring much less subjectivity than items on Facet 1, Facet 2 and Facet 3.

As with the PCL-R, a MANOVA was used to examine whether the clinicians reported that items on some HCR-20 scales required less subjective judgment than those on other scales. As with the PCL-R, subjectivity composite ratings were averaged for the items on each scale, separately for each clinician. The MANOVA showed that there was a significant difference in ratings for the three scales. Items on the Historical scale were seen as requiring much less subjectivity than items on the Clinical and Risk Management scales ($d = 3.05$ and 2.68, respectively). The effect size for the difference in subjectivity ratings between the Clinical and Risk Management scales was small ($d = 0.09$). Implications for measure development and refinement are discussed.

**Project and Scholarly Work**


**Summary**

The long-term predictive validity of the HCR-20, Static-99, SVR-20, and PCL-R was examined among 134 male sex offenders in Germany. Participants included contact and non-contact sex offenders, none of whom had an Axis I diagnosis. Exclusion criteria included being found to be not guilty by reasons of insanity, or with diminished capacity due to severe mental disorders; being dead; or having emigrated. Some inclusion criteria included having been imprisoned. The sample comprised three groups. The assessment only group consisted of 46 participants who were accused of sexual crimes and who were assessed for criminal responsibility in the forensic psychiatry department between 1975 and 1995. The treatment group consisted of 73 incarcerated sex offenders who underwent a two-year specialized psychosocial prison-based treatment program between 1972 and 1995. The treatment refuser and treatment drop-out group comprised the 15 individuals who did not complete this treatment program for a variety of reasons. Most participants were single (46.3%) and White Germans (94.2%). Participants’ mean age was 30.7 years.

Two individuals trained in the administration of the measures coded all cases. Assessments were completed retrospectively between 2001 and 2003 on the basis of file information. All measures were coded for the treatment and drop-out groups, but only the Static-99 was coded for the assessment group because of missing data. Recidivism data were obtained from the National Conviction Registry and were coded for the period following the completion of the assessment. Four types of re-offences were considered: (1) any non-sexual and non-contact criminal offence; (2) any non-sexual violent offence; (3) any non-contact “hands-off” sexual offence; and (4) any contact “hands-on” sexual offence. All analyses were calculated for the worst re-offence (i.e., type 4). The mean follow-up time was 9 years (range: 1 - 340 months). Time at risk ended at the first entry into the National Conviction Registry for any kind of recidivism.

None of the total scores on the four risk assessment measures was significantly better in predicting violent recidivism. AUC values for the total sample without the treatment drop-out group were as follows: HCR-20 total (AUC = .65, $p = .01$, SE = .05, 95% CI: .55-.75); Historical scale (AUC = .68, $p = .01$, SE = .05, 95% CI: .57-.78); Clinical scale (AUC = .58, $p = .16$, SE = .06, 95% CI: .47-.68); Risk Management (AUC = .48, $p = .68$, SE = .06, 95% CI: .37-.59); PCL-R total (AUC = .64, $p = .01$, SE = .05, 95% CI: .54-.74); SVR-20 total (AUC = .68, $p = .00$, SE = .05, 95% CI: .58-.78); Static-99 total (AUC = .72, $p = .00$, SE = .05, 95% CI: .62-.82). Values for the Static-99 for the total sample including the treatment drop-out group were: (AUC = .71, $p = .00$, SE = .05, 95% CI: .62-.80).

AUC values predicting all re-offences for the total sample excluding the treatment drop-out group were: HCR-20 total (0.67, $p < .01$); PCL-R total (0.65, $p = ns$); SVR-20 total (0.68, $p < .01$). The AUC value for the Static-99 for the total sample including the treatment drop-out group was 0.73 ($p < .001$).

AUC values predicting all non-contact sexual re-offences (which were included in the “all re-offences” category) for the total sample excluding the treatment drop-out group were: HCR-20 total (.41, $p = ns$); PCL-R total (.65, $p = ns$); SVR-20 total (.54, $p = ns$). The AUC value for the Static-99 for the total sample including the treatment drop-out group was .74 ($p < .05$).

AUC values predicting all contact sexual re-offences (which were included in the “all re-offences” and “all violent re-offences” categories) for the total sample excluding the
treatment drop-out group were: HCR-20 total (.67, \(p < .01\)); PCL-R total (.60, \(p = \text{ns}\)); SVR-20 total (.68, \(p < .01\)). The AUC value for the Static-99 for the total sample including the treatment drop-out group was .66 (\(p < .01\)).

Finally, the assessment and treatment groups were analyzed separately for predicting violent re-offending. In the assessment group, the accuracy of the Static-99 was numerically superior (AUC = .79, \(p = .00\)) than the other measures. In the treatment group, the Static-99 (AUC = .67, \(p = .028\)) performed better than the HCR-20 (AUC = .63, \(p = .07\)), PCL-R (AUC = .61, \(p = .10\)), and SVR-20 (AUC = .65, \(p = .03\)). The differences between instruments in the treatment group were not statistically significant.

---

**Project and Scholarly Work**


**Summary**

The present study was conducted to investigate the relationship between substance abuse and violence risk in individuals with psychopathic personality disorder. Male patients (\(N = 519\)) from forensic psychiatric clinics and correctional settings were assessed using the HCR-20 and PCL:SV. A subset of this group with psychopathy based on the PLC:SV (\(n = 138\)) were the focus of this study. Based on their ratings for item H5 (substance abuse problems), this subset was subsequently separated into those without a substance abuse problem (rating of 0), those with a substance abuse problem (rating of 1) and those with substance dependence (rating of 2). Based on these ratings, 23 participants had no abuse, 37 participants had a substance misuse problem, and 78 had substance dependence.

Differences between the ratings on the HCR-20 of these three groups of participants were then investigated. Kruskal Wallis H tests revealed that the groups differed on several items of the HCR-20: H3 (relationship instability), H8 (early maladjustment), H10 (prior supervision failure), C4 (impulsivity), R3 (lack of personal support), R4 (noncompliance with remediation attempts), and R5 (stress). Groups also differed in their H scale, R scale and total scores. For instance, total scores were 29.05, 30.24, and 33.33 for the no abuse, misuse, and dependence groups. It was also found that Part 1 of the PCL:SV was negatively correlated with substance use problems, and Part 2 was positively correlated with substance use problems.

The results are discussed in terms of the implications for substance abuse on the ratings of risk assessment tools.

---

**End of Mixed Setting**
Project and Scholarly Work


**Summary**

The aim of this study was to compare HCR-20 scores in two groups of adolescents; those who had known victims (family members and acquaintances) and those with unknown victims (strangers). All adolescents (N=104) admitted to the Adolescents Program of the Philippe-Pinel Institute of Montreal between February 1998 and April 2003 were interviewed. Their families or foster families were also interviewed systematically. The mean age at the time of admission was 16 years (SD = 1.4) and all participants were from Quebec, Canada. The most serious offences committed by the participants were grouped into 11 categories: homicide, aggravated assault, assault, other violence (such as robbery), arson, sexual assault, morality related, threats, public disorder, theft, or drug related. The French version of the HCR-20 was completed for each case but several items were modified for use with adolescents (H4 Employment Problems, H7 Psychopathy, H9 Personality Disorder and H10 Prior Supervision Failure). The family victims included victims who were the father, mother, brother, sister, or grandparents of the offender (N=15 adolescents) and the known group included victims who were the natural parent’s new spouse, half siblings, teachers, peers, or other persons known to the adolescent whether a close relationship existed or not (N=48 adolescents). The unknown group included victims who were strangers to the adolescent (N=41 adolescents).

The mean ranks (mr) of the global score of the HCR-20 differed between the two groups (mr = 48.6 for family/known victims, mr = 64.1 for unknown victims, z = -2.49, p = .013). Significant differences were also found on the H scale (z = -2.43, p = .015) and the R scale (z = -2.28, p = .022). Analysis of individual items on the HCR-20 differed between the two groups for H2 (z = -2.72, p = .006), H10 (z = -3.67, p = .001); C2 (z = -2.50, p = .013), R2 (p = .041) and R3 (z = -2.63, p = .009). For all analyses, the mean rank scores were higher for adolescents who victimized strangers, compared to those who victimized family/known victims.

Project and Scholarly Work


**Summary**

This study compared the HCR-20 and the SAVRY in a sample of 108 male juvenile offenders who were referred from court for inpatient psychiatric assessment (36 randomly selected who were nonrecidivists, 36 nonviolent recidivists, and 36 violent recidivists). This was a pseudo-prospective/retrospective follow-up study conducted from comprehensive youth justice, police, mental health, medical, and social-demographic files. The juvenile offenders were, on average, 15.3 at admission, and 25.1 at follow-up. Hence, this study evaluates the predictive validity of later adult violence of adolescent offenders. Follow-up national criminal records were used to code violence. Most participants were Caucasian.

Interrater reliability, based on a subset of 36 files, was good. ICC1 for Total, H, C, and R scores was .86, .88, .80, and .77. Pearson correlation coefficients between the HCR-20 Total Score, H, C, and R and violence were as follows: .46, .42, .35, and .44. Corresponding AUC values were .79, .76, .73, and .78.

Mean HCR-20 scores differed significantly between nonrecidivists, nonviolent recidivists, and violent recidivists. Total, H, C, and R scale scores across these three groups were as follows: 15.1 vs. 20.0 vs. 23.9; 6.1 vs. 7.8 vs. 9.8; 4.5 vs. 5.8 vs. 6.5; 4.5 vs. 6.5 vs. 7.6.

Comparison to the SAVRY showed that the HCR-20 produced somewhat larger correlations and AUC values, although there were no tests of significance between the two measures. Logistic regression analyses with all HCR-20 and SAVRY subscales showed that the HCR-20 Total, H, C, and R scale scores remained significant in the final model along with the SAVRY total score.
End of Juvenile Settings
Project and Scholarly Work


Summary

This article examined ways in which the VRAG and HCR-20 have been introduced as evidence in United States case law. Data was obtained through a legal database search of published and unpublished court cases in Federal and State jurisdictions across the US. A total of 46 cases in 12 different jurisdictions were obtained. California had the largest number of citations (15), followed by Minnesota (8), and Washington (8). The most common cited use of the VRAG and HCR-20 occurred in sexual commitment hearings. Other uses centered on violence risk assessments for potential parolees and for conditional release, as well as one case dealing with an adolescent transfer to adult court and one death penalty case. A summary of these cases is provided, along with legal and policy implications for using these risk instruments in various legal proceedings. The authors note two specific areas of concern: use in death penalty cases and in transferring a juvenile to adult court. Recommendations for ethical use in practice discussed.

End of Case Law Reviews
Non-Empirical Studies

Project and Scholarly Work


Summary

Violence risk assessments are conducted routinely in psychiatric and correctional settings. One method with demonstrated reliability and validity for assessing risk of violence is the structured professional judgment (SPJ) model. In this article, an overview of the SPJ model and a brief review of the empirical literature supporting its use is presented. The authors also provide a clinical case example to demonstrate the use of the HCR-20 with a psychiatric patient being considered for increased hospital privileges and discharge to the community. Recommendations for clinical practice using an SPJ tool when assessing risk of violence are presented.

---

Project and Scholarly Work


Summary

Violence risk assessments such as the HCR-20 and VRS-2 are increasingly being employed with female offenders. The PCL-R is not a violence risk assessment per se; however it is increasingly being employed with female offenders to inform assessment of risk. In this review, the applicability of violence risk assessment tools for female offenders is examined.

An overview of differing factors relevant to the assessment of male and female offenders suggests that pathways to offending differ across gender. Factors including abuse, substance use and mental health problems appear particularly characteristic of female pathways into offending behavior. Traditional assessments of risk employed with male offenders may not fully inform conceptualization of risk. There are also some variables including self-esteem, health, parenting difficulties, trauma, and forms of victimization that inform assessment of risk in female offenders that may not be covered in traditional risk assessments.

Research employing the HCR-20 and PCL-R with female populations is reviewed. With regards to the HCR-20, final risk judgments as opposed to total HCR-20 scores appear to be more predictive of risk for female offenders. This raises questions as to whether other factors besides those covered in the HCR-20 are employed in such a final judgment as HCR-20 scores have been found to predict violent outcomes in males more accurately than with females. Overall, it was found there is increasing levels of evidence supporting the use of the HCR-20 to inform conceptualization of risk with women when risk judgments as opposed to scores are employed, although further research is required regarding what other factors may further aid the HCR-20 when employed with this population. Research on use of the PCL-R with female samples has demonstrated links between levels of psychopathy and institutional violence, recidivism, and future violence. It seems, however, that the affective and interpersonal traits consistent with Factor 1 may be a better predictor of future offending than the anti-social and unstable lifestyle traits reflected in Factor 2. This is in conflict with the literature on male offenders although further research is required to explore whether this association between Factor 1 is also found with institutional aggression and future violence in female samples.

Overall, the research supports the use of the HCR-20 and PCL-R with female populations to inform conceptualization of risk of violent reoffending. Whether more complex assessments are required to capture women's risk of future violence are also explored.
End of Non-Empirical Studies
With regard to predicting violent recidivism, the HCR evidenced the largest mean correlation coefficient of .31 and the largest mean weighted $r$ of .28 (based on 11 effect sizes and a total sample size of 1395). For the LSI-R, the mean $r$ was .25 and the mean weighted $r$ was .28. For the PCL-R, the mean $r$ was .24 and the mean weighted $r$ was .27. For the VRAG, the mean $r$ was .27 and the mean weighted $r$ was .32.

Comparisons were also made between second generation tools (actuarial assessment instruments) and third generation tools (structured theory based assessment instruments). For the prediction of institutional violence, second generation instruments outperformed third generation instruments (mean weighted $r = .34$ and .20, respectively). For the prediction of violent recidivism, third generation instruments outperformed second generation instruments (mean weighted $r = .23$ and .18, respectively).

Next, the assessment tools were classified according to their inclusion of only static items, only dynamic items or a combination of both. For the prediction of institutional violence, static instruments yielded the largest effects, followed by combination instruments and dynamic instruments (mean weighted $r = .32$, .23, and .21, respectively). For the prediction of violent recidivism, dynamic instruments yielded the largest effects, followed by static instruments and combination instruments (mean weighted $r = .25$, .22 and .20, respectively).

Subsequently, risk assessments were classified according to how the information was gathered to complete the assessments: file review only, interview only, self-report, and combination of file review and interview. For the prediction of institutional violence, the mean weighted correlation coefficients ranked these approaches as follows: file review only ($r = .34$), combination of file review and interview ($r = .22$), self-report ($r = .16$), and interview only ($r = .14$). For the prediction of violent recidivism, the mean weighted correlation coefficients ranked these approaches as follows: combination of file review and interview ($r = .30$), file review only ($r = .26$), self-report ($r = .12$), and interview only ($r = .11$).

Several additional analyses are reported. The authors discussed the implications of these results in terms of the most appropriate instruments to use in professional practice.
Project and Scholarly Work


Summary

This meta-analysis consisted of a comprehensive examination of the Structured Professional Judgment (SPJ) model. This document contains an extensive review of the risk assessment field, guidelines for the selection of assessment tools, an introduction to the SPJ model, an overview of the various SPJ tools that are available for professional use, and a review of the existing research literature focusing on SPJ tools.

Studies were included in this meta-analysis if they contained data regarding an SPJ tool and an outcome of antisocial behaviour regardless of sample type or other sample characteristics. Nine online databases were searched for journal articles using multiple search terms. Next, reference sections of previous meta-analyses and other primary research were consulted. Several additional strategies were used to find unpublished research. This search resulted in 115 disseminations included in the meta-analysis, which represented 106 independent samples and 1881 effect sizes. Effect sizes were converted to AUCs when necessary. The majority of studies were from peer reviewed journals (61.1%) or conference presentations (15.9%) and were conducted in Europe (52.2%) or Canada (31.9%).

The mean age of participants across the studies was 31.14 (SD = 9.49). Most of the studies consisted of all male samples (52.2%). But many reported on mixed gender samples (32.4%). Base rates were as follows: violence (k = 149, M = 30.37%), physical violence (k = 109, M = 28.16%), sexual violence (k = 109, M = 19.83%) and other antisocial behaviour (k = 213, M = 43.03%).

Across the SPJ instruments examined the mean AUC was .66 and the mean weighted AUC was .65. With regard to the final risk judgments, the mean AUC was .68 and the mean weighted AUC was .68. With regard to the total score only, the mean AUC was .68 and the mean weighted AUC was .68. Including only one effect size per independent sample, the mean AUC was .70 and the mean weighted AUC was .68. Separating the outcomes into distinct types of antisocial behaviour and looking at the effects of total scores and final risk judgments the results were as follows: any antisocial behaviour (total score k = 93, AUCw = .68; final risk judgment k = 32, AUCw = .69), violence (total score k = 22, AUCw = .74; final risk judgment k = 32, AUCw = .71), physical violence (total score k = 40 AUCw = .68; final risk judgment k = 9, AUCw = .77), and sexual violence (total score k = 14, AUCw = .60; final risk judgment k = 9 AUCw = .65).

Focusing only upon the HCR-20, the largest effect size was found for the final risk judgments when the outcome was physical violence that included sexual violence (AUCw = .79). With regard to the prediction of any antisocial behaviour, the AUCw were .70, .69, .67, .65, and .65, for the final risk judgments, total score, H, C, and R scales, respectively. With regard to the prediction of violence, the AUCw were .76, .73, .70, .69, and .71, for the final risk judgments, total score, H, C, and R scales, respectively. With regard to the prediction of any physical, including sexual, violence, the AUCw were .79, .67, .66, .63, and .63, for the final risk judgments, total score, H, C, and R scales, respectively. With regard to the prediction of non-violent antisocial behaviour, the AUCw were .67, .70, .64, .68, and .68, for the final risk judgments, total score, H, C, and R scales, respectively.

The effect sizes yielded by different types of assessment approaches were also compared. SPJ instruments yielded the largest effect sizes for the final risk judgments and total scores both with mean AUC = .68 and AUCw = .68. The PCL was close behind with mean AUC = .65 and AUCw = .66. Actuarial instruments yielded slightly smaller effects with mean AUC = .61 and AUCw = .62. Unstructured clinical judgment yielded the smallest effects with mean AUC = .60 and AUCw = .59. Studies that compared an SPJ instrument directly to an actuarial instrument were examined separately. In all cases, the effect sizes were nearly identical with the SPJ instruments yielding slightly larger effects in some cases. However, no statistically significant differences were found.

Numerous additional results are reported. A section of this dissertation is spent investigating possible moderators of the relationship between SPJ instruments and antisocial outcomes. In brief, age, nationality, sample type (correctional, psychiatric, etc), setting (community, inpatient, etc), source of information used by raters, allegiance of researchers, or gender did not moderate the relationship. Of note, there was a trend for SPJ instruments to yield larger effects with female only samples compared to male only samples. The only moderator that was identified was the study design, with postdictive studies yielding larger effects than truly predictive studies.

Results are reported for additional SPJ instruments separately, as well as several actuarial instruments and the PCL. These results are not reported due to space limitations, as the current study is extensive (over 250 pages).
discussion section detailed the main findings and addressed clinical and research implications of the results.

### Project and Scholarly Work


### Summary

Psychopathy is one of the most researched violence risk factors and is included in numerous risk assessment instruments, including the HCR-20. This study questioned whether it was necessary to include psychopathy, as measured by the PCL-R or PCL:SV, in the HCR-20. The authors commented and supported theoretically and empirically the need to consider traits related to psychopathic and antisocial personality; however, the explicit inclusion of a PCL derivative was questioned. Furthermore, the impact of including or excluding the PCL in the HCR-20 on its predictive validity was a main focus. Various meta-analytic approaches were used to clarify these issues.

To begin with, data were gathered from published journals via electronic searches of eight databases using HCR-20 as the main search term, as well unpublished data was gathered from various sources. This sample consisted of 55 disseminations and 53 independent samples. The average sample size per dissemination was 171 (SD = 252).

This sample was used to determine the predictive validity of the full HCR-20 compared to the HCR-20 without H7 (psychopathy). With regard to the prediction of violence, the mean weighted AUC for the full HCR-20 was .69, whereas the mean weighted AUC was .70 for the HCR-20 without H7. With regard to the prediction of any antisocial behaviour, the mean weighted AUC for the full HCR-20 was .69, compared to .71 for the HCR-20 without H7. The Q statistic was not statistically significant for any of these analyses, indicating that the inclusion or exclusion of item H7 did not result in any change to the overall predictive validity.

Next, direct comparisons were made between the predictive the validity of the HCR-20 and the PCL-R in the 34 samples in which both tools were used. For the prediction of violence, within these samples, the mean weighted AUC for the HCR-20 was .69 and the mean weighted AUC for the PCL-R was .67. These weighted effect sizes did not differ significantly. Similar results were seen for the prediction of any antisocial behaviour.

Subsequently, five studies in which the HCR-20 was scored without using the PCL-R were examined. AUC values for the HCR-19 across these samples ranged from .62 to .81, with a mean of .72. This value is comparable to those previously reported. Overall, these analyses exhibited that the predictive validity of the HCR-20 is not dependent upon or overtly influenced by the inclusion or exclusion of item H7.

In the second part to this study, 16 raw datasets were gathered from researchers that included the HCR-20. These datasets represented 14 distinct samples. Of the 16 datasets, 9 were comprised of forensic psychiatric samples, 4 were correctional samples, and 3 were civil psychiatric samples. The samples ranged in size from 34 to 404 participants with an average of 150 (SD = 84). The majority of the samples consisted of males (approximately 80%) with a mean age of 35.54 years (SD = 4.59). The base rate of violence across these datasets was 32.8% (SD = 14.82).

With regard to the prediction of violence, the mean weighted AUC for the full HCR-20 total score was .73 and it remained at .73 when item H7 was removed. Item H7 alone resulted in a smaller mean weighted AUC of .61. Multivariate logistic regression was used to determine the independent effects of the HCR-20 and item H7. When The HCR-20 without H7 and item H7 alone were entered in a logistic regression analysis, the HCR-20 without H7 yielded a mean weighted $e^b$ of 1.04 indicating positive prediction of violence. Item H7 alone was unrelated to violence with a mean weighted $e^b$ of .99.

Next, the HCR-20 and PCL-R were compared directly using a subset of 7 datasets. The mean weighted AUC for the HCR-20 without H7 was .76, compared to .67 for the PCL-R. The incremental validity of these tools was investigated next. The HCR-20 without H7 was entered in the first block of a hierarchical logistic regression resulting in a significant model. The mean weighted $e^b$ for the HCR-20 without H7 was 1.21. When the PCL-R was entered in the second block, no significant improvement in overall model fit was seen. Moreover, the $e^b$ of the HCR-20 without H7 remained the same. These analyses were then conducted with the opposite order entry. When the HCR-20 without H7 was entered in the second block after the PCL-R, a significant improvement in the model was seen. Therefore, the HCR-20 evidenced incremental validity over the PCL-R, but the opposite was not true.

Numerous implications for the HCR-20 and clinical implications were discussed. Although the authors stressed that consideration of psychopathic and antisocial personality must be included in a comprehensive risk assessment, the statistical evidence indicated that the predictive validity of the HCR-20 was not dependent upon or influenced by the inclusion of the PCL-R. Alternative
methods for measuring and assessing psychopathic and antisocial traits are discussed throughout the paper.

See Also


Project and Scholarly Work


Summary

The present study investigated factors that were predictive of inpatient violence in forensic psychiatric units. Although many studies have investigated inpatient violence in civil psychiatric and correctional settings, there is much less research on forensic psychiatric samples. As such, the authors conducted this meta-analysis. Paper based and electronic literature searches were conducted to find studies that had investigated the perpetration of inpatient violence in forensic psychiatric settings. Only research within the last 15 years and including mostly male participants were included. Only factors that were included in at least three independent studies were investigated. From the various inclusion and exclusion criteria, 16 studies were included in the meta-analysis.

Factors that were entered into the analyses were age, scores on the BPRS, scores on the HCR-20, and scores on the PCL-R and PCL:SV. Based on 4 effect sizes (n = 243), a mean weighted r of .33 was found for the HCR-20 total score. Based on 6 effect sizes (n = 414), a mean weighted r of .19 was found for the H scale. Based on 5 effect sizes (n = 327), a mean weighted r of .35 was found for the C scale. The PCL-R, based on 3 effect sizes (n = 254), was also predictive of inpatient violence (mean weighted r = .21), as was the PCL:SV, based on 8 effect sizes (n = 827), with a mean weighted r of .26. Age was also found to be a modest predictor of inpatient violence. The BPRS total score was not predictive of inpatient violence.

Of all the factors that entered into the analysis, the largest effect sizes were found for the C scale and HCR-20 total score. It was concluded that the HCR-20, PCL-R, and PCL:SV are appropriate for assessing risk of inpatient violence in forensic psychiatric settings. These results were discussed in terms of their relation to the existing body of literature.

Project and Scholarly Work


Summary

Over the last 15 years, there has been a steady rise in the number of women committing violent crimes. Further, the number of female offenders in the criminal justice system is growing at a faster rate than that of their male counterparts. Compared to the research on male violence risk assessment, current and existing research on female violence risk assessment is sparse and limited. This meta-analytic study examined the overall effectiveness of violence risk assessments (HCR-20, PCL-R, LSI-R and VRAG) on females compared to males. Data for the study were obtained through searches of the following databases for articles published between 1990 and 2010: Academic Search Premier, Criminal Justice Periodical Index, EBSCOhost Research Databases, EBSCOhost Electronic Journals Services (EJS), Health and Psychosocial Instruments (HaPI), LexisNexis Academic, List of Online Journals, ProQuest Social Sciences CJ Periodicals, ProQuest Criminal Justice Journals, PsycARTICLES, PsycBOOKS, PsycCRITIQUES, PsycEXTRA, PsycINFO, Research Library, Social Science Research Network, and SociINDEX with Full Text. Unpublished manuscripts, dissertations, technical reports, and conference presentations were not included. Studies had to met the following inclusion criteria: (1) administration of violence risk assessments to female samples, with or without a matched sample of males for comparison, (2) sufficient statistical information allowing the calculation of the effect size statistic, (3) inclusion of adult female subjects, ages 18 or older, and (4) conducted with correctional, psychiatric, community, or custodial samples. The meta-analysis revealed that for men, the HCR-20 and the PCL-R had positive summary effect sizes, indicating that they have some predictive utility. The summary r was found to be .40 for the HCR-20 (p < .001) and .28 for the PCL-R (p < .001). For women, the HCR-20 (r = .31, p < .001) and LSI-R (r = .22, p < .001) were found to have nonzero effect sizes. However, the PCL-R (r = .10, p > .05) was not found to have a significant summary effect size. The author concluded that there is a severe lack of research that is available on violence risk assessments and their use and effectiveness with the female offender population.
Project and Scholarly Work


Summary

A meta-analytic review of published and unpublished research derived from psychiatric, correctional, adolescent, and mixed samples utilizing the HCR-20 (N = 6033) was conducted to obtain an overall estimate of the effectiveness of the scale, and to identify need for revision. Sixty samples were identified from 57 disseminations. The forensic psychiatric population contributed 3393 participants from 35 samples. The civil psychiatric setting accounted for 446 participants over four samples. For the correctional setting, there were 1363 non-disordered offender cases from 11 samples, and 368 mentally disordered offenders from five samples. There was a single youth sample (N = 108). Finally, there were 355 participants from 4 mixed samples.

Interrater reliability (IRR), averaged across 28 studies, was acceptable (Total score = .80; H = .88; C = .75; R = .75). Across the 36 studies that addressed predictive validity, HCR-20 total score and subscale scores were predictive of violence, with moderate to large effect sizes (.66 -.71). All reported average effect sizes are based on one effect size per study. Large effect sizes (.68 -.79) were observed across different types of clinical-legal contexts on the total HCR-20 and H scale. On the C scale, there were moderate to large effect sizes (.60 -.77). Small to high effect sizes were observed on the R scale (.48 -.78). Similarly, there were moderate to large effect sizes (.68 -.72) across different types of violence for HCR-20 total score and the C and R scales (.65 -.82; .64 -.72, respectively). The H scale scores had moderate effect sizes (.66 -.69). The AUC values for SPJ decisions for physical violence based on 4 samples (n = 725) were: .85 (de Vogel & de Ruiter, 2004); .70 (Fujii, Lickton, Tokioka, 2004); .78 (Douglas, Yeomans & Boer, 2005); and .74 (Douglas, Ogloff & Hart, 2003).

A number of moderator variables were examined. The context of violence outcome (i.e., institutional vs. community) also was examined as a moderator for mean AUC values. Moderate to large effect sizes were observed for institutional violence (Total = .71; H = .68; C = .73; R = .62), as well as for community violence (Total = .72; H = .71; C = .67; R = .69).

Moderate to large AUC values were observed among samples comprising only men (Total = .72; H = .70; C = .68; R = .70). Only a single effect size from a female sample was available (AUC for HCR-20 total = .62). Samples comprising both genders yielded moderate effect sizes (Total = .70; H = .68; C = .67; R = .63).

Samples were analyzed for method of coding the HCR-20. The instrument was coded either by review of file alone, or by a combination of file review and interview. For samples using the ‘file only’ approach, moderate to large AUC values were found (Total = .71; H = .68; C = .67; R = .66). For samples employing both file review and interview, large effect sizes were reported for the H (.73), C (.71) and Total (.73) scores. A moderate effect size was found for the R scale (.66).

Publication status seemed to have minimal impact: effect sizes for published (Total = .71; H = .68; C = .69; R = .65) and unpublished disseminations (Total = .71; H = .70; C = .66; R = .67) were of moderate to large magnitudes.

The country from which the samples were obtained was dichotomized as originating from either within or outside of North America. For North American samples, moderate effect sizes were observed (Total = .70; H = .70; C = .69; R = .69). Moderate to large AUC values were found for the remainder of the samples (Total = .71; H = .68; C = .66; R = .64).

Overall, results supported the robustness of the HCR-20 and will be used to guide recommendations for scale improvement.

Project and Scholarly Work


Summary

This paper presented the results from a systematic review and meta-analysis of studies of the predictive efficacy of the Historical, Clinical, and Risk Management-20 (HCR-20) for aggressive behavior in residential psychiatric facilities. Variations in efficacy were investigated based on aggression type (e.g., any, physical, verbal), HCR-20 scale used, and whether they were moderated by any clinical, demographic and methodological variables. Data for the study were obtained using several strategies. First, comprehensive terms were used to search seven electronic sources, including PsycINFO, CINahl, EMBASE, PubMed, MEDLINE, PsycLIT and SCOPUS. A comprehensive electronic search strategy was used to identify studies of the HCR-20 subscales. The HCR-20 total score and subscale scores were predictive of violence, with moderate to large effect sizes (.66 -.71). All reported average effect sizes are based on one effect size per study. Large effect sizes (.68 -.79) were observed across different types of clinical-legal contexts on the total HCR-20 and H scale. On the C scale, there were moderate to large effect sizes (.60 -.77). Small to high effect sizes were observed on the R scale (.48 -.78). Similarly, there were moderate to large effect sizes (.68 -.72) across different types of violence for HCR-20 total score and the C and R scales (.65 -.82; .64 -.72, respectively). The H scale scores had moderate effect sizes (.66 -.69). The AUC values for SPJ decisions for physical violence based on 4 samples (n = 725) were: .85 (de Vogel & de Ruiter, 2004); .70 (Fujii, Lickton, Tokioka, 2004); .78 (Douglas, Yeomans & Boer, 2005); and .74 (Douglas, Ogloff & Hart, 2003).

A number of moderator variables were examined. The context of violence outcome (i.e., institutional vs. community) also was examined as a moderator for mean AUC values. Moderate to large effect sizes were observed for institutional violence (Total = .71; H = .68; C = .73; R = .62), as well as for community violence (Total = .72; H = .71; C = .67; R = .69).
Mean weighted effect sizes (d_w) indicated that higher scores on the HCR-20 were predictive of various indices of inpatient aggression. In every comparison, mean weighted effect sizes for the dynamic (C and R) scales and the summary judgment were larger than the H scale. Further, with the exception of physical aggression against objects, the mean weighted effect sizes of the summary judgments were larger than all other scales. For the HCR-20 total score, C and R the large effect sizes, 0.93, 0.97 and 0.98, respectively, were obtained for verbal aggression. Smaller (though still ranging from moderate to large) effect sizes 0.42, 0.74 and 0.54, respectively were obtained for physical aggression against others. The largest effect size for H10, 0.42, was obtained for any inpatient aggression; whereas the smallest effect size, 0.3, was obtained for verbal aggression. For the HC15 composite, the largest effect size, 0.73 was obtained for physical aggression against objects. The smallest effect size, 0.47 was calculated for the category of any physical aggression. All effect sizes for the summary judgment were in the medium-large range. The largest effect size, 1.17 was found for any inpatient aggression and the smallest, 0.71 was for physical aggression against objects. It was not possible to calculate a weighted mean effect size for verbal aggression for the summary judgment as only one study predicted verbal aggression using this approach. Further analyses revealed that mean weighted effect sizes significantly different as a function of predictor scale, but not as a function of aggressive outcome. As such, subsequent analyses were conducted using the any inpatient aggression category for each predictor variable.

Overall, the HCR-20 had best predictive efficacy among samples containing higher proportions of patients with schizophrenia (HCR-20 total, H and R scales), women (HC15), and Caucasians (H). Predictive efficacy was reduced in studies containing higher proportions of patients with personality disorder. Few (n = 4) studies reported methodology and results sufficiently to ensure a transparently low risk of bias. The authors’ recommended that future research should aim to verify current findings using more heterogeneous samples and should report methodology with greater rigor.
Using a ranking system which collated findings of median AUC, PPV, NPV, and pooled DOR analyses, the authors found that the SAVRY, produced the highest overall rates of predictive validity, while the LSI-R, and the PCL-R, produced the lowest. Subgroup analysis and metaregression indicate that instruments produced higher rates of predictive validity in older and in predominantly White samples. The authors concluded that risk assessment procedures and guidelines by mental health services and criminal justice systems may need review in light of these findings.

See Also


Project and Scholarly Work


Summary

The study authors conducted a systematic review from 1990 to 2011 to identify and describe the psychometric properties and item content of measures aiming to predict outpatient violence risk in psychiatric patients, including schizophrenia. Risk assessment tools were identified using the search of online databases and Boolean keywords, by examining previous reviews, and through discussion with researchers in the field. Ten risk instruments designed to predict community violence in psychiatric patients (COVR, HCR-20, HKT-30, SAPROF, SORM, START, VRAG, V-RISK, VRS and the UK700) were identified. Studies investigating the predictive validity of these 10 tools were located using the same databases and time specifications above.

Information was collected on 11, 720 participants in 46 samples from 35 studies that investigated the ability of these 10 instruments to accurately predict community violence. The tools with the most samples were the HCR-20 (43.5%), VRAg (19.6%), and HKT-30 (10.9%). Further, the majority of the samples (65.2%) used an SPJ instrument. To examine item content and reporting characteristics, two of the study authors developed a checklist composed of 11 items pertaining to risk assessment tools organized into 3 scales (item content, validity, and reliability).

The item content of the 10 identified tools was classified into four categories: criminal history, sociodemographic, clinical, and neurobiological. The most common items on these scales were clinical factors, followed by sociodemographic characteristics, criminal history variables, and neurobiological factors. The clinical item that appeared most frequently was previous and/or current substance abuse (N tools = 9). The most common sociodemographic item was employment (N = 7). The most common criminal history variable was any previous and/or current violence (N = 6). Finally, only 1 neurobiological factor was identified: history of head injury with or without loss of consciousness (N = 1).

Each of the 10 tools was graded using the 11-item checklist developed for the study. Those instruments meeting the most checklist criteria included the HCR-20 (N criteria satisfied = 10), HKT-30 (N = 8), VRAG (N = 8), and VRS (N = 8). The most common criterion met was the inclusion of risk factors (N = 9). In relation to validity, it was the prospective measurement of predictive validity (N = 7), and on the reliability subscale, it was the testing of interrater reliability (N = 8). Less common were the inclusion of protective factors as item content (N = 5), the testing of divergent validity (N = 4), and the measurement of internal consistency (N = 1).

Independent predictive validity estimates were reported in 44 (95.7%) of the included samples. Performance indicators included the AUC (k cases = 43), sensitivity (k = 8), odds ratio (OR; k = 7), specificity (k = 6), positive predictive value (PPV; k = 4), negative predictive value (NPV; k = 4), product-moment correlation coefficient (r_p; k = 2), false positive rate (k = 1), and the number needed to detain (NND; k = 1). At least one AUC, OR, or r_p significantly above chance was reported for all of the included instruments. For 2 instruments (HCR-20 and VRAG) effect sizes were reported for different diagnostic groups. One interesting finding was that there appeared to be high levels of heterogeneity in these effect sizes across diagnostic groups. Only two studies that reported predictive validity estimates specifically in schizophrenic patients were found (median AUC = .69). When any diagnostic group was considered, mixed evidence or predictive validity was found, with median AUCs ranging from .62 to .85 depending on the population.

Convergent validity estimates were reported for 6 tools: the HCR-20, HKT-30, SAPROF, START, VRAg, and VRS. Evidence of significant divergent validity was found for the HCR-20, SAPROF, START, and VRS. Independent
reliability estimates were reported for 27 (58.7%) samples. At least 1 intraclass correlation coefficient above 0.75 or kappa coefficient (j) above 0.80, indicative of excellent interrater reliability was identified for the HCR-20, HKT-30, SAPROF, SORM, START, VRAG, V-RISK-10, and VRS. In addition, the HCR-20 was found to have a high level of internal consistency (α = .82).

The authors observed that a number of violence risk assessment tools exist that can be used to predict the likelihood of community violence in psychiatric patients; however, there were few studies that reported effect sizes specifically for patients with schizophrenia. The authors recommend that, when statistical power permits, studies should report psychometric properties for specific diagnostic groups. The authors concluded that there was large variation in item content between instruments, and further research was necessary to determine whether the inclusion of alternative factors could improve risk assessment.

**Project and Scholarly Work**


**Summary**

The current meta-analysis attempted to determine which risk assessment tools evidenced the best predictive ability with regard to violence. The authors based the current research question on an extensive theoretical backing reviewing several issues involved in violence risk assessment, including problems inherent to the procedure, the selection of appropriate tools, the definition of violence, the role of psychopathy in this endeavour, and a review of previous meta-analytic work. Subsequent to reviewing the previous meta-analyses, the authors highlighted a number of methodological flaws and misconceptions in this earlier work. Attempting to remedy these flaws, the authors chose nine widely used instruments to include in the present meta-analysis (the PCL-R, the PCL:SV, the HCR-20, the VRAG, the OGRS, the RM2000V, the LSI/LSI-R, the GSIR, and the VRS). They also attempted to reduce the between-individual sampling error by only including studies that compared the efficacy of more than one of these instruments on the same sample. Finally, they used multilevel regression models in order to test for random effects, compare the weighted effect sizes between instruments and correct for study features.

Studies were included if they met all of the following criteria: the study (a) assessed a sample with more than one of the risk assessment instruments, (b) included outcome data regarding violence, (c) reported sufficient statistical detail to calculate relevant effect sizes, and (d) was published or reported after 1999. With these criteria in mind, the authors searched three electronic databases (PsycINFO, Embase, and Medline) using several keywords. Specific journals known to include relevant studies were also searched. From the studies identified thus far, the reference lists were reviewed to identify any missing studies. Finally, reviewers of the present manuscript provided some additional studies. A total of 28 independent studies from 1999 to 2008 were included in the meta-analysis.

Studies included in the meta-analysis were carried out predominantly in the United Kingdom (k = 11) and Canada (k = 9), with others occurring in Sweden, Holland, United States, and Germany. Both truly prospective (k = 9) and quasi-prospective (k = 19) studies were included (reported as retrospective studies in the article; however, the description appears to indicate they are in fact quasi-prospective), with both community and institutional follow-up periods. Studies include samples of prisoners (k = 12), forensic psychiatric patients (k = 8) and offenders with mental disorders (k = 8). Participants had a mean age of 33.3 years. The majority of the samples were comprised of only males (k = 17), while 9 studies used mixed samples, and 2 studies used all female samples. The average follow-up period was 43.8 months, with an overall base rate of violence of 24.9%.

A Cohen’s d effect size was calculated from all studies, resulting in 174 effect sizes. Numerous study features were coded to determine any potential moderator effects, including participant mean age, gender makeup of the sample, study type, country, type of participant, definition of violence, and follow-up length. Several multilevel regression models were used with the PCL-R total score used as a reference because it was the mostly widely used instrument in all of the studies.

The first meta-regression, ignoring study heterogeneity, revealed that several instruments yielded larger effect sizes than the PCL-R, including the VRAG, the HCR-20, the PCL:SV, the OGRS, the GSIR, the RM2000V, and the VRS. As well, the HCR-20 H scale, the VRS static scale, the VRS dynamic scale, and PCL-R Factor 2 scores yielded larger effects than the PCL-R. In contrast, the PCL-R Factor 1, and the HCR-20 C scale and R scale yielded smaller effect sizes than the PCL-R. Subsequently, taking into account the heterogeneity between studies and the correlations between instruments within studies, another model revealed that only the HCR-20 total score had a larger
effect size than the PCL-R total score ($\chi^2(1) = 12.86, p = .0003$). Moreover, the PCL-R Factor 1 score had a smaller effect size than the PCL-R total score ($\chi^2(1) = 21.36, p < .0001$). No other differences were evident.

Further analysis attempted to account for sample characteristics that differed between studies. Results revealed that several features contributed to the effect sizes including the country of study, type of study (prospective versus quasi-prospective), length of follow-up and participant gender. After controlling for these study features, the effect sizes yielded for the HCR-20 was significantly larger than the PCL-R ($\chi^2(1) = 12.45, p = .0004$). The effect size yielded for the PCL-R Factor 1 was also still significantly smaller than the PCL-R ($\chi^2(1) = 20.89, p < .0001$). No other differences were found between the other instruments and the PCL-R. After taking into account these study features, the predictive validity of the risk instruments fell in the medium effect size range with AUCs between .56 and .71, with the majority falling between .65 and .69. With regard to the HCR-20, the AUCs, after controlling for these other features, were .71, .67, .66, and .66, for the total score, H, C and R scales, respectively.

With regards to study features more specifically, the studies conducted in the United States yielded smaller effect sizes compared to those conducted in Canada ($\chi^2(1) = 20.99, p < .001$). Prospective studies yielded larger effects than quasi-prospective studies ($\chi^2(1) = 4.82, p = .028$). Moreover, studies with longer follow-up period yielded larger effects sizes ($\chi^2(1) = 7.73, p = .0005$). Finally, studies with only female samples and mixed gender samples yielded larger effect sizes than all male samples ($\chi^2(1) = 5.68, p = .017$ and $\chi^2(1) = 9.03, p = .0038$, respectively). With regard to specific instruments, no gender differences were seen for the PCL-R, PCL:SV and HCR-20 total scores.

The main conclusion drawn from the data was that the HCR-20, for both genders, and the OGRS, for males only, yielded larger effect sizes than the PCL-R. The results are discussed in terms of the implications for each of the risk assessment instruments. Overall, the authors concluded that the selection of risk assessment instruments should be based on additional features of the instruments and not only on their predictive efficacy, for all the instruments essentially perform at the same level. Implications are discussed in terms of violence prediction, clinical applicability and violence management.
Scholarly Works that were Summarized in Section 2
(followed by the page number on which they appear)


Bauer, P., & Knörnschild, C. (2010, May). The ignored female minority: Do women have differentiated needs in the forensic setting? In R. Müller-Isberner (Chair), Forensic patients with special needs. Symposium conducted at the annual conference of the International Association of Forensic Mental Health Services, Vancouver, British Columbia, Canada. ................................................................. 68


Coid, J. (2009, June). All predictive instruments have the same predictive abilities. In J. Coid (Chair), Deconstructing risk assessment for violence. Symposium conducted at the annual conference of the International Association of Forensic Mental Health Services, Edinburgh, Scotland. .........................138


Assessment and Management: Bringing Science and Practice Closer Together. Sundsvall, Sweden, .......................... 90


Dolan, M., & Blattner, R. (2010). The utility of the Historical Clinical Risk -20 Scale as a predictor of outcomes in decisions to transfer patients from high to lower levels of security-A UK perspective. BMC Psychiatry (Open Access Journal) .......................................................... 81


Dolan, M., & Khawaja, A. (2004). The HCR-20 and post-discharge outcome in male patients discharged from medium security in the UK. Aggressive Behavior, 30, 469-483. .......................................................... 81


HCR-20 Review and Annotated Bibliography


HCR-20 Review and Annotated Bibliography


Henagulph, C.T. (2012). Risk and recovery group for offenders with mental disorders. *Psychiatric Services, 63*(1), 94-95. 103


Hilterman, E. & Chakhssi, F. (2002, March). *Prospective assessment of risk: comparing HCR-20, Behavioural Status Index (BSI) and Leave Risk Assessment (LRA)*. Paper presented at the Annual Conference of the International Association of Forensic Mental Health Services, Munich, Germany. 103


Petersen, K., Nicholls, T., Brink, J., & Webster, C. (2009, June). Convergent and divergent validity of the START with the PCL:SV and HCR-20. In T. Nicholls (Chair), Implementation and evaluation of START in civil and forensic psychiatric services. Symposium conducted at the annual conference of the International Association of Forensic Mental Health Services, Edinburgh, Scotland.

Petersen, K.L., Douglas, K., & Nicholls, T. (2012, March). Dynamic nature of risk assessed by the HCR-20 in a long-term civil psychiatric population experiencing...
HCR-20 Review and Annotated Bibliography


**Other Scholarly Works Cited in Section 2**


**Awaiting Copies of the Following Presentations**


presented at the 41st Annual Meeting of the American Academy of Psychiatry and the Law. Tucson, Arizona, USA.


Mullaney, R. (2012, June). Case mix, care, risk and constraint. In H. Kennedy (Symposium Moderator), The Use of Constraints in Forensic Mental Health Settings, Symposium presented at the annual conference of the International Association for Forensic Mental Health Services, Miami, Florida, USA.


Zareena Abidin, Mary Davoren, & Harry Kennedy. (2012, June). The complimentary role of risk assessment and assessment of programme completion and forensic recovery in transfer and release decision making. In H. Kennedy (Symposium Moderator), Discharge Criteria and length of stay for high, medium, and low secure facilities, Symposium presented at the annual conference of the International Association of Forensic Mental Health Services, Miami, Florida, USA.

Awaiting Copies of the Following Publications


## HCR-20\(^{\text{V1/V2}}\) Related Disseminations and Academic Commentary

### HCR-20\(^{\text{V1/V2}}\) Professional Manuals


### HCR-20\(^{\text{V1/V2}}\) Risk Factor Literature Review and Rationale


translated into Castellaño


translated into Catalan


translated into Dutch as:


translated into French as:


translated into German as:


translated into Norwegian as:


translated into Spanish as:


translated into Swedish as:


**Select HCR-20\textsuperscript{V1/V2} Published Peer Academic Commentary\textsuperscript{1}**


\textsuperscript{1} All of the disseminations in this section spend at least some time on the HCR-20. Some are devoted largely to the HCR-20, whereas others provide more circumscribed discussion.


### Related HCR-20\textsuperscript{V1/V2} Conference Presentations


Kullgren, G. (2001, April). Clinical (C5) and risk (R-5) scores according to the HCR-20 as related to treatment factors and criminal recidivism. Paper presented at the International conference of the International Association of Forensic Mental Health Services, Vancouver, British Columbia, Canada.


at the annual conference of the International Association for Forensic Mental Health Services, Miami, Florida, USA.

Related Professional Manuals

**Professional Manuals**


We invite any and all qualified persons to carry out research on the HCR-20. Our hope is to understand its psychometric properties as well as possible, and this requires empirical investigation at different sites and in various diverse samples.

We ask any persons who may have carried out research on the HCR-20 to please forward a copy of any presentations, manuscripts, or publications that emerge there from. Documents may be sent to the contact information provided on the following page. We aim to keep this annotated bibliography as current as possible, with as much existing research as possible.

Thank You!
Contact Information

The HCR-20 is copyrighted by the Mental Health, Law, and Policy Institute (Copyright Holder), at Simon Fraser University, and is distributed by (A) ProActive Resolutions, and (B) Psychological Assessment Resources. See below for contact information.

To contact Kevin Douglas at Simon Fraser University:
Kevin S. Douglas, LL.B., Ph.D.
Associate Professor
Scholar, Michael Smith Foundation for Health Research (2005-2010)
Department of Psychology
Simon Fraser University
8888 University Drive
Burnaby, British Columbia
Canada, V5A 1S6
Phone: 778-782-9562
Fax: 778-782-3427
Lab: 778-782-7239
Email: douglask@sfu.ca
Internet: http://kdouglas.wordpress.com/ or www.HCR-20.com

Guest Professor of Applied Criminology
Mid-Sweden University
Sundsvall, Sweden

To contact Psychological Assessment Resources, Inc.:
Internet: http://www4.parinc.com/

To contact ProActive Resolutions:
Internet: http://www.proactive-resolutions.com/

To contact the Mental Health, Law, and Policy Institute, Simon Fraser University:
Mental Health, Law, and Policy Institute
Simon Fraser University
8888 University Drive
Burnaby, British Columbia
Canada, V5A 1S6
Phone: 604-291-5868
Fax: 604-291-3427
Email: mhlpi@sfu.ca
Internet: http://www.sfu.ca/mhlpi