

AGAINST THE STREAM

The HCR-20 and violence risk assessment – will a peak of inflated expectations turn to a trough of disillusionment?

Edward Silva 

BJPsych Bulletin (2020) 44, 269–271, doi:10.1192/bjb.2020.14

Ashworth Hospital, Mersey Care NHS Foundation Trust, Liverpool, UK

Correspondence to Edward Silva (ed.silva@merseycare.nhs.uk)

First received 25 Oct 2019, final revision 17 Jan 2020, accepted 28 Jan 2020

© The Author 2020. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Summary The HCR-20 has taken on a life of its own. In forensic services it has been elevated from helpful aide-mémoire into a prophetic tool worthy of Nostradamus himself. Almost every outcome is interpreted through it. Despite the evidence of its limited utility, the difficulties of predicting rare events, the narrative fallacies and other heuristic biases it creates, and the massive opportunity costs it entails, commissioners and services alike mandate its use. Yet in routine practice the problems are not acknowledged, multiple conflicts of interest lie unobserved and other opportunities are neglected.

Declaration of interest E.S. is involved in the use of SPJ tools as part of the routine clinical care of detained patients.

Keywords Risk assessment; violence; forensic; HCR-20; SPJ.

Violence risk assessment is a core part of forensic psychiatry. It has evolved from an unstructured clinical and anecdotal exercise, through the use of actuarial tools and is now dominated by a variety of structured professional judgement (SPJ) instruments. Of these the Historical Clinical Risk Management-20 (HCR-20) is pre-eminent. It has itself evolved: first published in 1995, it is now in its third iteration. Initially it was used as an aide-mémoire to assist clinicians and others to systematically assess what were believed to be risk factors for violence across time: historical (ten items), clinical and risk management (five each). In 2001 further materials were added, including scenario planning.

The HCR-20 is the most widely used violence risk assessment tool in the world,¹ and in the UK it has become the ubiquitous gold standard for the risk assessment of violence in forensic services. NHS England commissioners of secure services for forensic patients mandate an HCR-20 assessment, updated every 6 months, even when there is no history of violence. As can be the case with many expert judgements,² any outcome can be seen through its lens. In cases of disaster, ‘the HCR-20 was completed incorrectly’, ‘the recommendations were not followed’, ‘it was not updated on time’ and, most seriously, ‘there was no HCR-20’. When there is success then the merits of the risk assessment and assessor are praised. Fearing blame in the event of failure, my psychologist colleagues spend dozens of hours reading through volumes of notes and the outputs are so long as to be unreadable. Explanations of previous violence are formulated, estimates of risk made and future risk scenarios hypothesised. The tool is over-relied on to guide

patient management through complex systems of care, a task it cannot achieve. Curiously, updates are frequently done after clinical decisions about management have been made. But the limitations are not acknowledged and they are legion. Some are relevant to violence risk in general and others specifically to SPJ tools and the HCR-20 in particular.

Limitations to SPJ tools

- There is no grade 1 randomised controlled trial (RCT) evidence for the effectiveness of SPJ tools in reducing violence; the only RCT tested the Short Term Assessment of Risk and Treatability (START) and gave a negative result.³
- Most items in structured risk assessment instruments, especially the Psychopathy Checklist –Revised (PCL-R), and many in the HCR-20 do not predict violence.⁴
- Random combinations of risk factors are as useful as those assembled in standardised instruments.⁵
- The HCR-20 ignores pertinent facts regarding the importance of adherence to specific drug treatments and risk.⁶
- The area under the curve (AUC) measure of utility bears very little relevance to use in clinical practice and ignores the difficulty of prediction when base rates are low.⁷ It is a concept rarely used in other areas of medical practice, where positive predictive value (PPV) is the usual measure.
- As with any attempt to predict rare events⁸ (p. 170) the PPV of the HCR-20, as with other risk tools, is poor

and it produces many more false- than true-positive findings.⁹

- High-quality negative evidence regarding the utility of multiple risk tools is not noticed, is refuted and as yet has had no impact on commissioners or services.^{9,10}
- Intellectual and financial conflicts of interest in the publications on various SPJ tools are not mentioned.¹¹ Those who submit research papers on the HCR-20 and other risk instruments rarely, if ever, declare an interest in receiving fees from training in its use. Yet it is a 'product', like a pharmaceutical agent, and one for which they stand to gain financially if it is promoted. Similar conflicts may exist for those who conduct serious adverse incident reviews recommending improved use of risk assessment if this is also a service they provide on a commercial basis.
- The narrative explanations of risk formulations and future risk scenarios are accepted. They are not seen as rhetorical devices requiring empirical validation, unlikely to be correct in systems too complex for analysis. To make sense of the world humans require stories that examine concrete events, ignoring chance and the things that did not happen. Any recent salient event is a candidate to become the kernel of a narrative explanation.¹²
- Narratives combined with recent or high-profile events feed heuristic biases, including representativeness, availability and, most important, affect.^{13,14} In forensic services our patients have often violated basic human norms: rape, incest, murder, mutilation and losses of control.¹⁵ At times we will be disgusted. This is rarely acknowledged and instead there is a serious risk that an emotionally driven sense of disgust¹⁴ will result in the immediate generation of opinions for which the supporting evidence is subsequently found, with risk assessment becoming confused with the assessment of outrage¹⁶ and becoming a moral exercise.¹⁷
- Whatever our organisations may tell us, it feels as if there is only punishment for failure and so an increasing tendency to risk aversion is inevitable.¹⁸
- The definition of violence used in the HCR-20 is so broad (including verbal threats) as to be meaningless in the services we work in.

The consequences of ignoring these limitations

Ignoring these difficulties is not just a failure of a tool. It has enormous consequences for patients, professionals, the public and those who pay for our services. The patients we care for face prolonged detention and the opportunity cost of professional time that could be spent delivering interventions. The patients we do not care for face delays in accessing care, often untreated and in inadequate facilities in prison. As professionals we become preoccupied with avoiding failure instead of achieving improvement and it often feels like the risk that is being managed is the risk to ourselves and to, or even from, our organisations. An explicit analysis of risk will be an important part of a patient's treatment, but in the context of deficiencies in treatment and access to care, an HCR-20 will not protect us, or our organisations, from litigation or public criticism. Instead of trying to determine what the prospective risk is given the facts and the base rates, we anticipate how failure will be perceived in

hindsight. Those that fund our services complain that too many are detained,¹⁹ while removing funding from objective research.²⁰ Inquires²¹ continue to recommend interventions that do not work – case management,²² risk assessment and community treatment orders²³ – and themselves can fuel narrative fallacies.²⁴ Through our overvalued ideas regarding risk assessment, forensic services are left caring for a tiny percentage of mentally disordered offenders, who we dare not part company with, and at vast expense.²⁵

What can we do?

The argument is not that risk assessments should be abandoned, only that we should be much more circumspect about their power, utility and explanatory value, and recognise how narratives may mislead as well as explain. This is now the position in the related field of suicide risk assessment. In stark contrast to the requirements for secure services and the use of the HCR-20, the National Institute for Health and Care Excellence (NICE) advice is: 'Do not use risk assessment tools and scales to predict future suicide or repetition of self-harm'²⁶ (p. 8), for the simple reason that we cannot stratify risk using the tools available. The information they provide regarding the likelihood of the outcomes we are really concerned about is of no practical use.²⁷ But it is very hard for systems to change and for professionals to give up their sincerely held beliefs. This is the case throughout medicine. It takes an average of 17 years to translate research findings into practice.²⁸ Although short structured assessments would be helpful, our attempts to stratify risk of violence are not useful and should be abandoned, as should narrative explanations of the past and hypothesising future scenarios. It is not particularly useful to say that a man who has been violent in the past might be violent in future if intoxicated, threatened, feeling disrespected or aggrieved, lost to follow-up, non-adherent to anti-psychotic or mood stabilising medication and in contact with a vulnerable potential victim.

Some hope that technology will provide a solution. But it took the resources of Deep Mind's artificial intelligence (AI) capabilities, combined with a vast sample of over 700 000 patients, to develop a system to predict the highly specific outcome of acute kidney injury within the tight window of 48 h in highly monitored in-patient environments.²⁹ So why do we think that we can predict violent behaviour over timescales of weeks, let alone months or years, on the basis of human analysis, or that in future AI will be able to make longer-term predictions about far more complex human behaviours? Even if such analytic systems are developed, it is questionable whether clinicians, patients or the legal system would accept them. It is likely that highly discriminatory variables would be key factors in AI algorithms – gender, age, ethnicity, residence in a high crime area, peer group criminality – and there would be fears that the scenarios of *The Minority Report* would emerge.³⁰ Instead the approach adopted by NICE regarding suicide and self-harm should be taken, with the emphasis on the delivery of effective treatments, ensuring services are adequately resourced and developing better habits regarding quality.³¹

A hint of change?

A quick search using Google Trends shows that online interest in the HCR-20 has fallen dramatically, from a peak in September 2007 to date. The Gartner Hype Cycle,³² with its phases of a technology trigger, a peak of inflated expectations, a trough of disillusionment, a slope of enlightenment and then a final plateau of productivity, is held as an example of the boom, bust and then stabilisation of new technologies. But perhaps this is what is happening already?

About the author

Edward Silva is a consultant forensic psychiatrist at Ashworth Hospital, Mersey Care NHS Foundation Trust, Liverpool, UK. He has worked in secure services since 1998 and has been involved in the use of SPJ tools throughout as part of the routine clinical care of many detained patients.

References

- Douglas KS, Hart SD, Webster CD, Belfrage H. *HCR-20V3: Assessing Risk of Violence – User Guide*. Mental Health, Law, and Policy Institute, Simon Fraser University, 2013.
- Tetlock PE. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton University Press, 2005.
- Troquete NAC, van den Brink RHS, Beintema H, Mulder T, van Os TWDP, Schoevers RA, et al. Risk assessment and shared care planning in out-patient forensic psychiatry: cluster randomised controlled trial. *Br J Psychiatry* 2013; **202**: 365–71.
- Coid JW, Yang M, Ullrich S, Zhang T, Sizmur S, Farrington D, et al. Most items in structured risk assessment instruments do not predict violence. *J Forensic Psychiatry Psychol* 2011; **22**: 3–21.
- Kroner DG, Mills JF, Reddon JR. A coffee can, factor analysis, and prediction of antisocial behavior: the structure of criminal risk. *Int J Law Psychiatry* 2005; **28**: 360–74.
- Fazel S, Zetterqvist J, Larsson H, Långström N, Lichtenstein P. Antipsychotics, mood stabilisers, and risk of violent crime. *Lancet* 2014; **384**: 1206–14.
- Nielsen O, Large M. Rates of homicide during the first episode of psychosis and after treatment: a systematic review and meta-analysis. *Schizophr Bull* 2010; **36**: 702–12.
- Stone JV. *Bayes' Rule: A Tutorial Introduction to Bayesian Analysis*. Seibel Press, 2013.
- Fazel S, Singh JP, Doll H, Grann M. Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: systematic review and meta-analysis. *BMJ* 2012; **345**: e4692.
- Tully J. HCR-20 shows poor field validity in clinical forensic psychiatry settings. *Evid Based Ment Health* 2017; **20**: 95–6.
- Singh JP, Grann M, Fazel S. Authorship bias in violence risk assessment? A systematic review and meta-analysis. *PLoS One* 2013; **8**(9): e72484.
- Kahneman D. *Thinking, Fast and Slow*. Penguin, 2011.
- Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. *Science* 1974; **185**: 1124.
- Slovic P, Peters E. Risk perception and affect. *Curr Dir Psychol Sci* 2006; **15**: 322–5.
- Brown DE. Human universals and their implications. In *Being Humans: Anthropological Universality and Particularity in Transdisciplinary Perspectives* (ed N Roughley): 156–74. De Gruyter, 2013.
- Sandman PM. Hazard versus outrage in the public perception of risk. In *Effective Risk Communication: The Role and Responsibility of Government and Nongovernment Organizations* (eds VT Covello, DB McCallum, MT Pavlova): 45–9. Springer, 1989.
- Haidt J. The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychol Rev* 2001; **108**: 814–34.
- Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica* 1979; **47**: 263–91.
- Keown P, Murphy H, McKenna D, McKinnon I. Changes in the use of the Mental Health Act 1983 in England 1984/85 to 2015/16. *Br J Psychiatry* 2018; **213**: 595–9.
- Healthcare Quality Improvement Partnership. *The National Confidential Inquiry into Suicide and Safety in Mental Health. Annual Report: England, Northern Ireland, Scotland, Wales. October 2018*. University of Manchester, 2018.
- Crichton JHM. A review of published independent inquiries in England into psychiatric patient homicide, 1995–2010. *J Forensic Psychiatry Psychol* 2011; **22**: 761–89.
- Burns T, Creed F, Fahy T, Thompson S, Tyrer P. Intensive versus standard case management for severe psychotic illness: a randomised trial. *Lancet* 1999; **353**: 2185–9.
- Burns T, Rugkåsa J, Molodynski A, Dawson J, Yeeles K, Vazquez-Montes M, et al. Community treatment orders for patients with psychosis (OCTET): a randomised controlled trial. *Lancet* 2013; **381**: 1627–33.
- Chiswick D. The falling shadow: a psychiatrist's view. *J Forensic Psychiatry* 1995; **6**: 594–600.
- Wilson S, James D, Forrester A. The medium-secure project and criminal justice mental health. *Lancet* 2011; **378**: 110–1.
- National Institute for Health and Care Excellence. *Self-Harm in over 8s: Longer-Term Management [Clinical Guideline CG 133]*. 2011.
- Large MM, Ryan CJ, Carter G, Kapur N. Can we usefully stratify patients according to suicide risk? *BMJ* 2017; **359**: j4627.
- Morris ZS, Wooding S, Grant J. The answer is 17 years, what is the question: understanding time lags in translational research. *J R Soc Med* 2011; **104**: 510–20.
- Tomašev N, Glorot X, Rae JW, Zielinski M, Askham H, Saraiva A, et al. A clinically applicable approach to continuous prediction of future acute kidney injury. *Nature* 2019; **572**: 116–9.
- Jochelson R, Gacek J, Menzie L, Kramar K, Doerksen M. *Criminal Law and Precrime: Legal Studies in Canadian Punishment and Surveillance in Anticipation of Criminal Guilt*. Taylor & Francis, 2017.
- Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat A-HS, Dellinger EP, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2009; **360**: 491–9.
- Gartner. *Gartner Hype Cycle: interpreting technology hype*. Gartner Inc, 2016 (<https://www.gartner.com/en/research/methodologies/gartner-hype-cycle>).

