Understanding Clinicians’ Use of Cues When Assessing the Future Risk of Violence: A Clinical Judgement Analysis in the Psychiatric Setting

Barbara Brown1,2 and Tim Rakow1*
1 University of Essex, Colchester, UK
2 James Paget University Hospital, Gorleston, UK

Research is sparse on how clinicians’ judgement informs their violence risk assessments. Yet, determining preferences for which risk factors are used, and how they are weighted and combined, is important to understanding such assessments. This study investigated clinicians’ use of static and dynamic cues when assessing risk in individual patients and for dynamic cues considered in the recent and distant past. Clinicians provided three violence risk assessments for 41 separate hypothetical cases of hospitalized patients, each defined by eight cues (e.g., psychopathy and past violence severity/frequency). A clinical judgement analysis, using regression analysis of judgements for multiple cases, created linear models reflecting the major influences on each individual clinician’s judgement. Risk assessments could be successfully predicted by between one and four cues, and there was close agreement between different clinicians’ models regarding which cues were relevant for a given assessment. However, which cues were used varied between assessments: history of recent violence predicted assessments of in-hospital risk, whereas violence in the distant past predicted the assessed risk in the community. Crucially, several factors included in actuarial/structured risk assessment tools had little influence on clinicians’ assessments. Our findings point to the adaptivity in clinicians’ violence risk assessments, with a preference for relying on information consistent with the setting for which the assessment applies. The implication is that clinicians are open to using different structured assessment tools for different kinds of risk assessment, although they may seek greater flexibility in their assessments than some structured risk assessment tools afford (e.g., discounting static risk factors). Copyright © 2015 John Wiley & Sons, Ltd.

Key Practitioner Message:
• Across three separate violence risk assessments, clinicians’ risk assessments were more strongly influenced by dynamic cues that can vary over time (e.g., level of violence) than by static cues that are fixed for a given individual (e.g., a diagnosis of psychopathy).
• The variation in the factors affecting risk assessments for different settings (i.e., in hospital versus in the community) was greater than the variability between clinicians for such judgements.
• The findings imply a preference for risk assessment strategies that offer flexibility: either using different risk assessment tools for different purposes and settings or employing a single tool that allows for different inputs into the risk assessment depending upon the nature of the assessment.
• The appropriateness of these clinical intuitions about violence risk that are implied by our findings warrants further investigation.

Keywords: Violence Risk Assessment, Clinical Reasoning, Judgement Analysis, Clinical Prediction, Risk Factors, Dynamic Risk Assessment

INTRODUCTION
Assessing the potential for violence is an important element of forensic and clinical practice. In forensic work, such risk assessments inform sentencing and probation decisions. In mental health, violence risk assessments contribute to treatment and patient management decisions, including the choice of treatment environment (e.g., high-security units versus low-security units versus community treatment). A key goal in most risk assessments is to support risk management: to protect those who come into contact with a potentially violent individual and to contribute to effective treatment and/or management of...
that individual. And it follows that if mistakes are made when assessing risk, then clinical staff, community members or other patients may be put in harm’s way, and the standard of care may be diminished.

The merits of different approaches to assessing the risk of violence have been widely discussed. Early debates contrasted unstructured clinical judgement, which relies on a practitioner’s knowledge and expertise to identify and respond to risk factors, with actuarial assessment, which algorithmically combines previously identified risk factors into a risk score (Dawes et al., 1989; Doyle & Dolan, 2002). A third approach—structured professional judgement (SPJ)—broadens the focus of risk assessment beyond predicting violence to managing the risk of violent behaviour (Campbell et al., 2009; Douglas & Kropp, 2002; Douglas & Skeem, 2005; Heilbrun et al., 2009). This approach relies on structured tools (e.g., checklists) but allows for professional judgement in identifying additional patient or contextual factors and in weighting, combining or responding to risk factors. Thus, SPJ represents something of a ‘middle way’, and in this vein, the UK Department of Health (2007) guidelines reflect a compromise between clinical and actuarial risk assessments by promoting evidence-based practice that maintains clinical flexibility.

Numerous studies, including some within the domain of violence risk assessment, suggest that actuarial approaches to risk prediction are more accurate than unstructured clinical judgement (Dawes et al., 1989; Gardner et al., 1996; Grove & Meehl, 1996). Structured tools (grounded in the SPJ approach) have been observed to predict violence as well as actuarial tools (e.g., for intimate partner violence; Messing & Thaller, 2013) and show the same kind of advantage over unstructured judgement as that found for actuarial approaches (Campbell et al., 2009). This advantage for structured/actuarial tools over unstructured judgement is generally assumed to be because they ensure that important risk factors are considered and appropriately weighted and—particularly for actuarial assessment—because they simplify the difficult cognitive task of reliably combining multiple factors (Einhorn, 1974). Indeed, good practice guidelines provided by the UK National Institute for Clinical Excellence (2005) and the UK Department of Health (2007) recognize pure clinical judgement in risk assessments as incomplete and prone to bias and therefore promote the use of structured risk assessments such as Historical Clinical Risk 20 (HCR-20), Violence Risk Scale, Violence Risk Appraisal Guide (VRAG) and Classification of Violence Risk. While the items included and the format of the assessment differ between these tools, a factor analysis of the items incorporated in such assessments by Kroner, Mills and Reddon (2005) produced similar dimensions across the tools: criminal history, antisocial lifestyle, psychopathy and substance abuse/mental health issues. However, formal assessments derived from actuarial tools are not always embraced by clinicians, who sometimes view them as impractical (Quinsey et al., 2005), replacing their expertise (Slovic et al., 2000) or a poor representation of their professional reality (Doyle & Dolan, 2002).

These tools rely on different types of risk factor (i.e., cues/variables that predict violence). Some approaches to risk assessment distinguish between static and dynamic cues, according to whether a feature can vary over time for a given individual (Blackburn, 2000)—and this distinction is important for the study we report here. Thus, variables such as sex, age of first conviction or personality disorders that are resistant to change (e.g., psychopathy) are typically regarded fixed features of the individual and so classed as static cues. However, over time, a person may exhibit differing levels of violence, may express different attitudes or may change their employment status—all of which are therefore regarded as dynamic cues. Note that for dynamic cues, one can consider current or past values of these variables and, if past, that these might be recent or distal values. This past/present distinction is important in other taxonomies of risk factors. For example, Monahan and Steadman (1994) distinguish between cues that are historical (e.g., previous violence and childhood adjustment), dispositional (e.g., cognitive, emotional or social tendencies), clinical (e.g., impulsiveness and delusions) or contextual (environmental/situational features somewhat external to the individual; e.g., their number of friends). Unfortunately, these alternative cue taxonomies do not map neatly onto one other (e.g., oftentimes, historical features are also static cues, but historical and static are not synonyms). Moreover, cues do not attract unique classifications; for instance, Monahan and Steadman (1994) describe psychopathy as a clinical cue, whereas the HCR-20 structured assessment tool classifies it as historical rather than clinical.

Evidence that the cues used by clinicians in making violence risk assessment may differ from the ones indicated by structured tools that came from Elbogen, Calkins-Mercado, Scalora and Tomkins (2002). They used surveys to examine the importance that clinicians assign to the factors included in actuarial and structured risk assessments as well as to other variables. Their results indicated that clinicians recognize actuarial/structured assessments as an important element of violence assessment. Nevertheless, clinicians regard dynamic cues as more relevant in assessing violence (including some behaviours that are omitted from structured tools). Clinicians’ responses suggested that their judgement relies heavily on clinical cues.

\(^1\)Department of Health guidelines are published to provide advice and guidance for health professionals across the UK to ensure that patients receive care of the best quality. These do not replace professional expertise but, rather, seek to describe and encourage best practice.
such as current symptoms of mental illness, responsiveness to treatment and psychosocial adjustment. More distal historical factors (often static cues), such as educational or social history (e.g., childhood maladjustment), work history and psychopathy, have been shown to predict violence; however, clinicians viewed these as less relevant for violence risk assessments.\(^2\)

These clinicians’ self-reports of their own risk assessment are interesting and potentially important for clinical practice. However, it seems that people often have limited insight into the factors that affect their choices (Berry & Broadbent, 1984; Nisbett & Wilson, 1977) or have difficulty in accurately articulating their judgement processes (Hastie & Dawes, 2010; Holzworth, 2001; Hogge, 2001). Nonetheless, the findings of Elbogen et al. (2002) suggest an intriguing possibility that clinicians’ violence risk assessments focus on recent behaviours at the expense of those from the more distant past—even when evidence suggests that such distal historical factors are predictive of violence (e.g., Loeb et al., 2005).

This paper examines clinical judgement in violence risk assessment, and responds to calls for more research into the information that clinicians use when making risk assessments (Elbogen et al., 2002; Skeem & Monahan, 2011). In order to address the shortcomings of self-report survey measures, we employ a quantitative methodology—clinical judgement analysis (Holzworth, 2001)—which, by analysing a clinician’s judgements, indicates which cues are most influential in those judgements. This allows us to estimate the relative impact of static versus dynamic cues on violence risk assessments, and for dynamic cues to compare the impact of recent versus distal behaviours (e.g., violence severity in the near or far past).

Even though current best practice recommends the use of structured and actuarial approaches, there are several reasons why it is important to better understand the role of clinical judgement in violence risk assessment. First, despite advice to the contrary, Higgins, Watts, Bindman, Slade and Thornicroft (2005) found that a sizeable minority of UK mental health institutions do not provide psychiatrists with structured instruments for violence risk assessment, and in other instances, training in the use of structured instruments is limited. If structured risk assessment is not uniformly employed, there is pragmatic value to understanding which factors are the dominant influences on clinical reasoning about violence risk. Second, even when actuarial risk scores are computed, clinical judgement plays a critical role. Practitioner judgements may be used as inputs into an actuarial tool (Einhorn, 1974; Menzies & Webster, 1995), or judgement may be formally combined with the output of an actuarial tool (Blatberg & Hoch, 1990). Less formally, risk scores may be used—alongside clinical judgements—as one of several inputs into risk assessments. Thus, understanding how clinicians respond to different risk factors could be key to appreciating when and how practitioners will augment actuarial risk assessments with additional factors or when they might discount, or even ignore, the outputs of actuarial tools. Third, and perhaps most crucially for current practice in violence risk assessment, SPJ provides for an explicit role for clinical judgement. Understanding clinicians’ preferences for utilizing different predictors when assessing risk is therefore key to understanding which additional factors are more likely to be introduced, or which factors will be weighted heavily, in a structured risk assessment. Finally, there is a range of actuarial and structured judgement tools available, which differ according to the particular collection of risk factors that they use. Although clinicians may sometimes use more than one tool, there is inevitably an element of selection when tools are adopted in clinical practice. Therefore, understanding which kinds of factor clinicians believe to be most predictive of violent behaviour may improve our understanding of clinicians’ preferences for selecting one assessment tool over another.

Our chosen technique for furthering our understanding of the role of clinician judgement in risk assessment is clinical judgement analysis, which models the judgement of an individual judge (e.g., risk assessor). Using judgements (e.g., risk assessments) for multiple cases, where several variables or ‘cues’ (e.g., risk factors) take different values from case to case, this technique obtains a model of the judge/assessor providing those judgements. This is most commonly performed by regressing judgement onto multiple cues, with the resulting linear model providing a description of the assessor’s ‘policy’ for making judgements. The standardized regression coefficients (β-weights) for each cue provide a guide to the relative weight given to that cue in the judgement process. Individual differences in cue use can be explored by comparing policies (linear models) from different assessors. Where the true predictiveness of a cue can be estimated (e.g., the strength of correlation between that risk factor and future violence from research), this can be compared with the corresponding β-weight from an assessor’s policy. Linear models derived in this way successfully predict future clinical judgement (Wigton, 2001), even when only three or four significant cues are used (Hastie & Dawes, 2010). Similarly—often contrary to the self-reports of the assessors themselves (Konecni & Ebbesen, 1984)—many professional judgements can be successfully represented as a simple process involving a few variables, which are combined without regard for configurial or non-linear relations.

Our study takes an approach that others have used: constructing a set of hypothetical cases for our participants to

---

\(^2\)Research findings are mixed with respect to whether clinical factors or historical factors best predict violence. This issue is explored in the Discussion section.
assess. Prior research suggests that assessments are similar for actual cases and hypothetical cases (Denig & Rethans, 1996; Kirwan et al., 1983). Hypothetical cases increase the efficiency of the research process by creating replicable scenarios that can be presented to many assessors, enabling assessment of judgement consistency both within and between individuals (Patterson et al., 2008). Additionally, the correlation between cues can be minimized to increase the statistical power to detect influences on judgement. Importantly, while hypothetical cases presented as written summaries lose some of the information available in clinical practice, they do retain an important element of ecological validity for the research: clinicians often make judgements on the basis of documented data (e.g., patient notes), much of which represent the opinion or observation of other professionals (e.g., nursing reports and daily care reports). Clinical judgement analysis is often described (perhaps unusually for a quantitative method) as ‘idiographic’—its primary aim is to understand the individual rather than the group. As such, it is important to have sufficient data points per participant (i.e., to present enough cases) to adequately model each clinician. Inevitably, practical limitations (e.g., our participants’ time) constrain us to present fewer cases, and therefore to consider fewer cues to judgement, than we would have wished to. Nonetheless, in keeping with our rationale set out above, we examine several examples of static and dynamic cues. For dynamic cues, we consider behaviours that are assessed within two periods (in the recent or more distant past).

METHODS

Participants

The target group for this study were clinicians having an active role in violence risk assessment (e.g., clinical psychologists, psychiatrists, social workers and senior nurses). We therefore approached potential participants with experience of conducting risk assessments and/or of using them to manage risk in psychiatric settings. Nine clinicians participated: one clinical psychologist, two nurses and one social worker (all working in one of two private psychiatric hospitals), three doctoral-level clinical psychology trainees (all with at least 3 years’ clinical experience), a psychological well-being practitioner and a trainee social worker (both with prior experience in a psychiatric unit/hospital). Thus, the various roles fulfilled by different members of a multidisciplinary team that would typically be involved in making and responding to violence risk assessments were well represented in our sample. Clinical experience ranged from 0 to 16 years, and only one participant had no prior experience of conducting risk assessments. A further four individuals agreed to participate but did not return the questionnaire within the time scale of the study.

Design and Materials

Participants received a booklet containing 41 hypothetical cases. Each case was described through a series of details of past events and behaviours (see Appendix 1 for an example; see Box 1 for the case ‘template’). Each case was defined according to eight cues, the levels of which varied from case to case (Table 1). The choice of cues was guided by Dolan and Dole’s (2000) analysis to reflect factors utilized across several risk assessment tools (e.g., HCR-20, VRAG and some key cues explicitly identified as important by clinicians (Elbogen et al., 2002). Cue names and their level labels followed naming conventions used in clinical practice and in assessment tools (e.g., severity of violence: mild, moderate or severe). To avoid repetition of precise details that might strike participants as unrepresentative, the exact values for the two levels for length of current presentation and date of first violent event were varied within narrow ranges (Table 1). Patient’s sex and age were fixed for all cases: as male and in the narrow range of 31–35 years (randomly assigned within this range).

The cases comprised a ‘test set’ of 32 cases, and a further ‘holdout set’ of nine cases (inter-mixed) to allow for cross-validation of the linear models. This number of cases reflected the practicalities of recruiting participants: we judged that including more cases would push the task completion time beyond the 60–75 min and that it was reasonable to ask of our participants. The number of cues per case then reflected the guidance of Cooksey (1996, p. 128) who recommends a minimum of 30 cases when eight uncorrelated cues are examined via clinical judgement analysis. Cue values for each case and the number of cases required were determined according

---

3Reflecting an idiographic approach (and, presumably, also the difficulty of recruiting experts as participants in time-consuming research), clinical judgement analyses are often small-N studies. For example, Kirwan et al. (1983) obtained judgement policies for rheumatoid arthritis prognoses from two clinicians, and Einhorn (1974) examined judgement policies for three pathologists. This shortcoming, which also applies to our study with \( n = 9 \), necessarily limits what can be concluded from any one study. Thus, the better summaries of this body of research (e.g., Hastie & Dawes, 2010; Slovic & Lichtenstein, 1971) have a meta-analytic flavour and focus on those findings that are consistent across multiple studies.

4Psychological well-being practitioners are trained to deliver psychological therapy to patients with depression and/or anxiety and work within the UK National Health Service (‘Improving Access to Psychological Therapies’ service).
to a fractional factorial design generated using SPSS ORTHOPLAN (IBM Corp.). This ensured an efficient design with uncorrelated cues within the test set, which reduces ambiguity in interpreting regression estimates and ensures standard errors for the standardized regressions coefficients (β) of each cue, which are close to 0.1 for the number of cases presented in the test set (Stewart, 1988). Cue values were ‘translated’ into the appropriate text descriptions by one author (B. B.) and ‘back-translated’ by the other (T. R.) to check for errors.

The cases were presented in a booklet in a single random order (as determined by SPSS ORTHOPLAN). Three risk assessment questions were presented for each case (Appendix 1). Two questions focussed on the probability of a patient exhibiting violent behaviours in the next 6 months: (a) if he is discharged to the community and (b) if he remains in the current hospital placement. The third question required assessing the level of danger the patient posed in the current hospital. These two separate questions posed for in-hospital risk probed two slightly different conceptions of risk. Probability of violence refers clearly to the chance of a violent episode and makes no reference to its likely severity or potential for harm. Level of danger allows for a broader conception of risk, which may prompt assessors to incorporate beliefs about the potential for harm should it occur. Thus, assessing danger might be broadly equivalent to assessing expected (dis)utility (from a decision theoretic perspective) or to a single combined assessment of the impact and the probability of an adverse event (from a risk management perspective). We expected that participants would differ more in their precise interpretation of the term ‘danger’ than would be so for ‘probability’ (Dowie, 1999), although we anticipated that judgements for the more inclusive term—danger—might better reflect clinician’s priority for action (see Cahill & Rakow, 2012; Slovic, 1987).

Participants responded to each question by placing a mark on the (10-cm) line of a Visual Analogue Scale (VAS) beneath each question. Each VAS represented a continuous 0–10 scale from low to high probability/danger. Additional categorical descriptors (‘low’, ‘medium’ and ‘high’) appeared above the VAS as a reference scale. Such categorical representation of risk frequencies is familiar to clinicians (Slovic, et al., 2000) and avoids presenting a scale that is not used naturally by clinicians (e.g. percentage estimates), which can distort responses.

Prior to data collection, to ensure that the task instructions were clear and that SPSS ORTHOPLAN had not generated any cases with implausible combinations of cue values, a pilot participant with some clinical experience completed the entire task.

### Procedure

Participants received the booklet of 41 study cases, which began with an instruction page (approximately 450 words) and ended with a brief demographic questionnaire requesting details of profession, sex and clinical experience (overall length and length and degree of experience with violence risk assessment).

The initial instructions informed participants of the focus of the project (clinical judgement) and identified the institution conducting the research. Participants were informed that the cases were hypothetical but designed to be realistic and were asked to imagine that the cases were to be reviewed at a Care Programme Approach (multidisciplinary team) meeting in a psychiatric hospital. The clinicians were asked to respond as they would in their everyday practice if presented with this case information and were instructed how to complete a VAS. The instructions stressed the voluntary nature of their participation, their anonymity in the analysis/reporting of results and that their clinical performance/training was not being assessed by the research. Participants were also informed that they could complete the booklet over ‘one or a few’ sessions. Participants were asked to complete the task within 2 weeks, and most participants accepted an offer to be contacted within a week to have a further opportunity to ask questions (researcher contact...
details were also provided). The time taken by participants to return the questionnaire ranged from 4 days to 3 weeks.

Research Ethics

The study received ethical approval from the University of Essex (UK) prior to commencement. Permission to approach clinicians employed at psychiatric hospitals (trainees) was obtained from hospital management (the course director) before doing so. Participants gave written consent before providing their responses. Participants were offered an opportunity to receive individual feedback regarding their judgement policies; several took up this offer and received this feedback once the data had been analysed.

RESULTS

In keeping with standard methods for clinical judgement analysis (Cooksey, 1996; Hogge, 2001), multiple linear regression was used to establish each participant’s cue utilization for each of the three risk assessment questions (i.e., separate analyses for each participant for each question) using only the data for the 32 cases in the test set. Therefore, three sets of regressions were performed using the eight cues as predictors (Table 1); the dependent variable for each set was the judgement (0–10 continuous scale) provided in response to one of the three risk assessment questions:

1. Probability of violent behaviour in the next 6 months if the patient is discharged to the community.
2. Probability of violent behaviour in the next 6 months if the patient remains at his current hospital.
3. The level of danger that the patient poses if he remains in the current hospital.

Where brevity is required, we label these assessment questions as ‘community risk probability’, ‘in-hospital risk probability’ and ‘in-hospital danger’, respectively. The results of the regressions are summarized in Tables 2–4. Each row reports the data for a single participant, starting with the more experienced clinicians in the upper rows.

Model Fit

The first two columns of data in Tables 2–4 provide the summary of the overall model fit for the regressions (represented by $R^2$ and adjusted $R^2$). These linear models accounted for a substantial proportion of the variance in participants’ judgement, ranging from 42% to 88% of the variance accounted for across all three of the risk assessment questions. This suggests that the linear model detected regularities in individual judgement patterns, although it did so to varying degrees across individuals and risk assessment questions. Risk assessments were more predictable (on the basis of cue values) for assessment Question 2 (in-hospital risk probability) and Question 3 (in-hospital danger) than for Question 1 (community risk probability): the mean (median) $R^2$ values were 0.60 (0.62) for Question 1, 0.71 (0.71) for Question 2 and 0.69 (0.70) for Question 3. Similarly, there was some variation across participants in the predictability (reliability) of judgement as measured by these models: mean $R^2$ values (from the three assessment questions) ranged from 0.53 for Assessor F and 0.58 for Assessor G to 0.85 for Assessor H. The mean $R^2$ values for the remaining six assessors clustered between 0.64 and 0.69.

Model Validation

Further insight into the regularities within each individual’s judgements comes from a comparison of predicted judgements derived from linear models ‘built’ from the test set data and actual judgements made for the holdout set of nine cases. Linear models were constructed for each judgement and for each individual risk assessor using only the significant cues from the regression analyses: each regression being rerun with only these significant cues to obtain the appropriate regression coefficients (shown in square brackets in Tables 2–4). A predicted judgement was then obtained for each of the holdout cases by substituting the cue values for each case into these linear models (with between one and four cues). The correlations between predicted and actual judgements are shown in the final two columns of Tables 2–4 (represented by $r$ and $r^2$). These provide further evidence of the ability of linear models based on a limited number of cues to capture a substantial proportion of the variance in judgement, although, again, there is some variation in ‘predictability’ between individuals and (to a lesser extent) between different risk assessment questions. The ability of the linear models to predict risk assessments was slightly higher for assessment Question 3 (in-hospital danger) than for Question 1 (community risk probability) or Question 2 (in-hospital risk probability): the mean (median) $r^2$ values were 0.53 (0.49) for Question 1, 0.52 (0.48) for Question 2 and 0.59 (0.57) for Question 3. The ‘least predictable’ assessor (Assessor E) according to this model fit criterion, with a mean $r^2$ value of 0.39, was not one of the two assessors who had been deemed least predictable according to their $R^2$ from the regression analysis of the test set (above). However, Assessor H was again the ‘most predictable’ individual with a mean $r^2$ value of 0.82. The mean $r^2$ for the other seven participants ranged from 0.47 to 0.68.
Table 2. Linear models of judgement for the probability of violent behaviour in the next 6 months if the patient is discharged to the community (Question 1)

<table>
<thead>
<tr>
<th>Assessor (Yrs exp.)</th>
<th>Overall fit of linear model</th>
<th>Standardized regression coefficients (β-weights) for individual cue use [unstandardized regression coefficients for linear model, significant cues plus constant only]‡</th>
<th>Model correlation with holdout set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
<td>Adjusted $R^2$</td>
<td>Psychopathy</td>
</tr>
<tr>
<td>A</td>
<td>0.620**</td>
<td>0.49</td>
<td>0.237</td>
</tr>
<tr>
<td>(10–19)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.620**</td>
<td>0.490</td>
<td>0.227</td>
</tr>
<tr>
<td>(4–9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.620**</td>
<td>0.490</td>
<td>−0.061</td>
</tr>
<tr>
<td>(4–9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>0.520*</td>
<td>0.350</td>
<td>0.183</td>
</tr>
<tr>
<td>(1–3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>0.420</td>
<td>0.210</td>
<td>0.126</td>
</tr>
<tr>
<td>(1–3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>0.480*</td>
<td>0.300</td>
<td>0.207</td>
</tr>
<tr>
<td>(1–3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>0.650**</td>
<td>0.520</td>
<td>−0.169</td>
</tr>
<tr>
<td>(&lt;1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>0.850***</td>
<td>0.790</td>
<td>0.058</td>
</tr>
<tr>
<td>(&lt;1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>0.620**</td>
<td>0.480</td>
<td>0.314*</td>
</tr>
<tr>
<td>(0)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Yrs exp. = years of experience of making violence risk assessment where participants indicated the same level of experience; the participant with longer overall clinical experience is listed first.

‡Referring to the most recent presentation of violent behaviour prior to the current admission to the hospital.

†Unstandardized coefficients are used for the linear models used to generate predicted judgements. For example, for Assessor H: expected judged probability of violent behaviour = 3.21* (Frequency of violence pre-admission) + 0.95*(Severity of violence pre-admission) – 2.75. Therefore, for this participant, each change of level for frequency of violence is associated with 3.21 increment on the 0–10 risk assessment scale.

*p < 0.05.

**p < 0.01.

***p < 0.001.
<table>
<thead>
<tr>
<th>Assessor (Yrs exp.)</th>
<th>Overall fit of linear model</th>
<th>Standardized regression coefficients (β-weights) for individual cue use [unstandardized regression coefficients for linear model, significant cues plus constant only]</th>
<th>Model correlation with holdout set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Length of current presentation</td>
<td>Frequency of violence pre-admission*</td>
</tr>
<tr>
<td>A (10–19)</td>
<td>0.650**</td>
<td>0.530</td>
<td>0.133</td>
</tr>
<tr>
<td>B (4–9)</td>
<td>0.730***</td>
<td>0.630</td>
<td>0.217</td>
</tr>
<tr>
<td>C (4–9)</td>
<td>0.710***</td>
<td>0.600</td>
<td>−0.147</td>
</tr>
<tr>
<td>D (1–3)</td>
<td>0.680**</td>
<td>0.560</td>
<td>0.120</td>
</tr>
<tr>
<td>E (1–3)</td>
<td>0.810***</td>
<td>0.740</td>
<td>−0.063</td>
</tr>
<tr>
<td>F (1–3)</td>
<td>0.640**</td>
<td>0.510</td>
<td>0.031</td>
</tr>
<tr>
<td>G (&lt;1)</td>
<td>0.570**</td>
<td>0.420</td>
<td>0.043</td>
</tr>
<tr>
<td>H (&lt;1)</td>
<td>0.830***</td>
<td>0.780</td>
<td>0.090</td>
</tr>
<tr>
<td>I (0)</td>
<td>0.730***</td>
<td>0.630</td>
<td>0.035</td>
</tr>
</tbody>
</table>

Yrs exp. = years of experience of making violence risk assessment where participants indicated the same level of experience; the participant with longer overall clinical experience is listed first.

*Referring to the most recent presentation of violent behaviour prior to the current admission to the hospital.

*p < 0.05.

**p < 0.01.

***p < 0.001.
### Table 4. Linear models of judgement for the level of danger posed if the patient remains in his current hospital (Question 3)

<table>
<thead>
<tr>
<th>Assessor (Yrs exp.)</th>
<th>$R^2$</th>
<th>Adjusted $R^2$</th>
<th>Psychopathy</th>
<th>Length of current presentation</th>
<th>Frequency of violence pre-admission$^*$</th>
<th>Severity of violence pre-admission$^*$</th>
<th>Behavioural problems at school</th>
<th>Date of first violent offence</th>
<th>Frequency of in-hospital violence</th>
<th>Severity of in-hospital violence</th>
<th>Constant</th>
<th>Model correlation with holdout set</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (10–19)</td>
<td>0.760***</td>
<td>0.680</td>
<td>0.557****</td>
<td>1.862</td>
<td>0.112</td>
<td>0.069</td>
<td>0.243*</td>
<td>−0.075</td>
<td>0.105</td>
<td>0.326**</td>
<td>0.504***</td>
<td>[−0.375]</td>
</tr>
<tr>
<td>B (4–9)</td>
<td>0.570**</td>
<td>0.420</td>
<td>0.305*</td>
<td>2.500</td>
<td>0.152</td>
<td>0.000</td>
<td>0.000</td>
<td>0.076</td>
<td>−0.152</td>
<td>0.506**</td>
<td>0.414**</td>
<td>[−8.267]</td>
</tr>
<tr>
<td>C (4–9)</td>
<td>0.750***</td>
<td>0.660</td>
<td>−0.107</td>
<td>0.498</td>
<td>0.000</td>
<td>0.097</td>
<td>0.032</td>
<td>0.214</td>
<td>0.000</td>
<td>0.162</td>
<td>0.808***</td>
<td>[−1.761]</td>
</tr>
<tr>
<td>D (1–3)</td>
<td>0.880***</td>
<td>0.840</td>
<td>0.119</td>
<td>0.498</td>
<td>0.085</td>
<td>0.371***</td>
<td>0.108</td>
<td>0.060</td>
<td>−0.062</td>
<td>0.239**</td>
<td>0.791***</td>
<td>[−2.196]</td>
</tr>
<tr>
<td>E (1–3)</td>
<td>0.680***</td>
<td>0.570</td>
<td>−0.066</td>
<td>0.570</td>
<td>0.066</td>
<td>−0.099</td>
<td>−0.099</td>
<td>0.066</td>
<td>−0.066</td>
<td>0.218</td>
<td>0.774***</td>
<td>[2.741]</td>
</tr>
<tr>
<td>F (1–3)</td>
<td>0.480*</td>
<td>0.290</td>
<td>−0.011</td>
<td>0.510</td>
<td>0.143</td>
<td>−0.026</td>
<td>0.181</td>
<td>0.305</td>
<td>0.032</td>
<td>0.461**</td>
<td>0.338*</td>
<td>[−4.162]</td>
</tr>
<tr>
<td>G (&lt;1)</td>
<td>0.510*</td>
<td>0.330</td>
<td>0.079</td>
<td>0.860</td>
<td>0.113</td>
<td>−0.041</td>
<td>0.087</td>
<td>0.162</td>
<td>0.089</td>
<td>0.660***</td>
<td>0.822***</td>
<td>[−2.191]</td>
</tr>
<tr>
<td>H (&lt;1)</td>
<td>0.860***</td>
<td>0.810</td>
<td>−0.001</td>
<td>0.860</td>
<td>0.009</td>
<td>−0.046</td>
<td>0.191*</td>
<td>0.001</td>
<td>−0.092</td>
<td>0.372**</td>
<td>0.723***</td>
<td>[−5.053]</td>
</tr>
<tr>
<td>I (0)</td>
<td>0.700***</td>
<td>0.590</td>
<td>0.081</td>
<td>0.590</td>
<td>0.024</td>
<td>−0.005</td>
<td>−0.003</td>
<td>0.073</td>
<td>0.248*</td>
<td>0.315*</td>
<td>0.723***</td>
<td>[2.502]</td>
</tr>
</tbody>
</table>

Yrs exp. = years of experience of making violence risk assessment where participants indicated the same level of experience; the participant with longer overall clinical experience is listed first.

*Referring to the most recent presentation of violent behaviour prior to the current admission to the hospital.

* $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$. 

---

**Violence Risk Assessment**

Copyright © 2015 John Wiley & Sons, Ltd.
Cue Utilization

The central columns of Tables 2–4 summarize each participant’s cue utilization for each risk assessment question, as inferred from the regression analyses. The codes for the cue values were such that positive regression coefficients reflect risk assessments that are consistent with the literature on risk factors for violence and their instantiation in risk assessment tools (e.g., greater risk: in the presence of psychopathy than in its absence, in the presence of behavioural problems at school or with an earlier rather than later date of first violent offence). The participants’ risk assessments were consistent with the literature: the great majority of the regression coefficients in Tables 2–4 were positive. Where negative coefficients did occur, these were almost always close to zero—indicative of minimal reliance on a cue, as opposed to using the cue in the reverse of the expected direction.

The regression coefficients for the statistically significant predictors of each participant’s risk assessments—which usually represent effects of medium size or greater (Cohen, 1992)—are shown in bold type in Tables 2–4. This serves to highlight two features of the data. First, that the frequency and severity of violence are the dominant influences on participants’ violence risk assessments—although the precise use of these cues varies between the three risk assessment questions (e.g., whether pre-admission or in-hospital violence is more important). Second, it highlights that participants’ risk assessments in this study can be characterized as reliant on a relatively small number of cues: the number of significant cues for a given set of judgements varied from one to four, depending upon the individual and the assessment question. This is not to say that participants never used more than four pieces of information when making their risk assessments—of course, the statistical significance of a cue reflects an arbitrary cut-off that is partly constrained by the statistical power associated with each risk assessment question (with 32 test cases per question). Nonetheless, the fact that model validation $r^2$ values were only slightly below the model fit $R^2$ values and were similar to the adjusted $R^2$ values for the model fits suggests that risk assessments can usually be reduced to (i.e., modelled as) a process involving two or three cues without substantial loss of information.

Table 2 shows that assessments of the probability of violence if the patient is discharged (risk assessment Question 1) relied heavily on the frequency of violence pre-admission: for the majority of participants, more frequent presentation of violent behaviours pre-admission substantially increased the assessed probability of violence in the community. Additionally, some participants also gave significant weight to the in-hospital frequency of violence or the severity of previous violence (pre-admission or in-hospital). With one exception (Assessor D), the regression analysis of participants’ assessments for this question revealed two or three significant cues.

Table 3 shows that the in-hospital frequency of violence is the dominant predictor of participants’ assessment of the probability of violent behaviour if the patient remains in his current hospital (risk assessment Question 2). For six participants, this was the only significant predictor of their assessment; the remaining three participants had two or three significant cues for this assessment question. Thus, more frequent presentation of violent behaviours in hospital (in the recent past) increased the assessed probability of violence in the hospital (in the near future)—and, in general, other factors had little bearing on this assessment.

Table 4 shows that the in-hospital frequency and severity of violence were the key influences on the level of danger posed by the patient if he remains in his current hospital (risk assessment Question 3). Every participant had at least one of these two cues as a significant predictor of his or her judgement, and six had both as significant predictors. This means that with increasing severity and/or frequency of violent behaviours in hospital, the level of dangerousness is perceived to increase; the tendency is for severity to take precedence over frequency in this assessment (i.e., higher $\beta$’s for severity than for frequency). The range for the number of significant cues per individual assessor was greater for this assessment question than for the other questions: three participants had only one significant cue, whereas three participants had three significant cues and one (Assessor A) had four cues that were significant predictors of his or her judgement.

Following from Elbogen et al. (2002), one goal of our research was to explore the relative weight that clinicians give to recent behaviours and to more distant-past behaviours when making violence risk assessments. Our design permitted a formal analysis of this with respect to the frequency and severity of violence, which were described in our cases for the recent past (in-hospital) and more distant past (pre-admission). To effect this analysis, we used the beta weights for frequency and severity shown in Tables 2–4 as scores to reflect the weight given to each cue. These were subjected to a repeated-measure ANOVA with three factors: assessment question (1 versus 2 versus 3), time of behavioural cue (‘distal’ pre-admission versus ‘recent’ in-hospital) and measure of violence (frequency versus severity). This three-way ANOVA with cue weight ($\beta$) as the dependent variable revealed a main effect of time, $F(1,8) = 27.7$, $p = 0.001$, $\eta^2 = 0.78$. On average, $\beta$-weights were higher for recent behaviours than distal ones ($M_{\text{recent}} = 0.38$ versus $M_{\text{distal}} = 0.15$). However, this effect was qualified by a significant question-by-time interaction, $F(2,16) = 28.0$, $p < 0.001$, $\eta^2 = 0.78$. Recent cues were more influential than distal ones for assessment Question 2 (in-hospital risk probability), $\Delta M_p = 0.43$, $t(8) = 18.0$, $p < 0.001$, and for Question 3 (in-hospital danger),
DISCUSSION

We used clinical judgement analysis to investigate the individual judgement policies of clinicians when making three separate violence risk assessments and identified several recurring patterns in these judgement policies. Cues representing recent or potentially ongoing (directly observable) behaviours were generally given more weight than cues relating to past events or to fixed features of the individual. Clinicians' risk assessments could be successfully modelled by, and predicted from, a handful of cues combined via simple linear rules. It is important to acknowledge that other factors (e.g., cues such as the patient’s gender, de Vogel & de Ruiter, 2005) or processes (e.g., non-linear cue use), which have been omitted from our investigation, might also be important when assessing risk. Nonetheless, our data point to some potentially important conclusions regarding cue utilization in violence risk assessment, how this varies across different types of violence risk assessment and the implications of this variation for clinical practice.

Several factors that have previously been shown to predict violent behaviour in psychiatric patients (e.g., psychopathy, Skeem & Mulvey, 2001) seemed to be given a relatively low priority by the clinicians in our study. For example, the presence/absence of psychopathy and the date of first violent offence6 were rarely included in the assessors’ judgement policies (Tables 2–4). Our participants were consistent in utilizing the frequency of violent behaviours in their risk assessments, and—to a somewhat lesser degree—the severity of those violent episodes. These findings chime with those of Elbogen et al. (2002), whose participants identified physical aggression while in care and the history of violent behaviour as the most relevant risk factors. While our participants varied in their level of experience and performed a variety of clinical roles, our small sample size precludes drawing conclusions about the effects of experience on clinician’s judgement in risk assessment or the differences in judgement associated with different clinical roles (e.g., nurses and physicians) or separate subspecialties (e.g., psychiatry and clinical psychology). While previous research has often found little variation in judgement policies for risk assessment by level of expertise or clinician specialty (e.g., Cheyne et al., 2012; Elbogen et al., 2002), one might expect that such differences could occur in violence risk assessment (e.g., reflecting some differences in the experiences, or goals, of different clinicians). Therefore, a larger study that used clinical judgement analysis to investigate the (potential) effects of expertise, clinical role and clinical subspecialty on violence risk assessment could be of considerable value.

For two of our three risk assessment questions, recent violent behaviours (in hospital) were prioritized over past violence (prior to admission). This may reflect a bias in human information processing to favour recent data when making judgements (Elstein, 1999). It is, of course, a reasonable assumption that, when predicting behaviour in the immediate future, behaviour in the recent past provides a better guide than behaviour in the more distant past. Indeed, this is often shown to be the case (e.g., in skilled performance, Henry & Hulin, 1987; in educational attainment, Humphreys, 1968), and it is also true that in dynamic environments, it is sometimes better to ignore the distant past and focus on recent events when making decisions (Rakow & Miler, 2009). However, retrospective assessments of behaviour are often insensitive to duration (Kahneman & Frederick, 2002). Therefore, what a simple distinction between the recent and distant past may overlook is that records of behaviour from the distant past may often reflect a larger sample of events—which may be a better guide to underlying influences or habitual patterns of behaviour than recent observations that might be heavily dependent upon current, but ephemeral, circumstances (e.g., contextual factors; Monahan & Steadman, 1994). Indeed, this was the case in our scenarios: recent violence was based on 6 months of observation, whereas distant-past violence was based on a sample of data ranging from 8 months to 7 years of accumulated observations (depending on the length of presentation of the current episode).

[6]The three-way interaction was also significant, \(F(2,16) = 10.9, p = 0.001, \eta^2 = 0.58\), although as none of our main research questions bore on this effect, we did not explore this further (Greenwald, 1993). None of the remaining main effects or interactions was significant in this ANOVA (all \(p > 0.1\)).

[7]Moreover, one participant appeared to weight the first offence cue in the opposite direction to that expected on the basis of actuarial risk assessments: HCR-20 indicates young age of first violent offence as the relevant factor in predicting violent behaviours (Doyle & Dolan, 2002).
There is, however, an alternative interpretation of the apparent priority of recent observations of violence over more distant ones. Note that the setting or location of those observations differed, as well as their timing. For judgements concerning future violence in the community (risk assessment Question 1), long-past behaviours outside of the hospital setting received greater weight than those observed recently in the hospital setting. Thus, our data are also consistent with the possibility that clinicians rely more heavily on behaviours previously observed in the setting to which the risk assessment applies. Recall that Questions 2 and 3 both required assessments of in-hospital risk and that for these assessments (recent) in-hospital behaviours were weighted more heavily. Therefore, perhaps contrary to the common tendency to favour dispositional attributions of behaviour over situational ones (Jones, 1979), it seems that our participants may have selected cues according to the context in which behaviours previously occurred.

Whatever the precise reason behind the differential cue use across separate risk assessment questions, our findings point to an important feature of our participants’ risk assessments: their cue use was configural in the sense that the weight given to certain cues was contingent upon which outcome was being assessed.7 Thus, participants’ assessments reflected an assumption that different kinds of violence risk assessment (in this case, for different settings) require different strategies for risk assessment. This has important practical implications for those involved in managing the risk of violence and is reflected in recent research on violence risk assessment.

Recent meta-analyses of violence prediction via actuarial and structured tools by Yang, Wong and Coid (2010) and by Campbell et al. (2009) provide some analysis relevant to this assumption. While Yang et al. found few differences in predictive accuracy across nine violence risk assessment tools, they did find that these tools predicted some outcomes more effectively than others. Broadly defined criminal recidivism was more effectively predicted than other violent/aggressive outcomes (e.g., narrowly defined recidivism and aggression within an institution). Importantly, Yang et al. highlighted that different assessment tools were designed for different purposes and that some may be more suited to particular functions than others. However, due to model-fitting constraints, they were unable to formally assess whether some tools were superior for certain outcomes. Likewise, in a rare direct comparison of violence prediction for different settings (hospital, community or prison), Menzies and Webster (1995) also note that prediction is more successful in some settings than others (and distinctly problematic in some). From their meta-analysis, Campbell et al. (2009) report that actuarial instruments dominated by static risk factors offered superior prediction of institutional violence, whereas structured measures that incorporate dynamic risk factors were better for predicting violent recidivism in the community (although these differences between locations were generally small). Certainly, one practical implication of our participants’ propensity to use different cues for different risk assessments is that it implies an openness to use SPJ tools flexibly (depending on the kind of assessment) or to adopt different assessment tools (incorporating different sets of risk factors) for assessing violence risk in different settings. The implication is that clinicians accept that one size does not fit all when it comes to risk assessment tools.

Further evidence that different contexts might require distinct risk assessment strategies can be gleaned by comparing the relative contribution of clinical and historical factors to the prediction of violence in different settings. Douglas, Ogloff, Nicholls and Grant (1999) found that historical factors (i.e., HCR-H) were more predictive than clinical ones (HCR-C) for predicting post-discharge violence. In contrast, McNeil et al. (2003) found that the HCR-C Scale was superior to the HCR-H Scale for predicting in-hospital violence, and suggested that assessments of short-term risk may require different tools for assessments of chronic violence risk. However, perhaps inevitably (given the difficulties in predicting violent behaviour), not all studies confirm this pattern of results. Gardner et al. (1996) found clinical history to be at least as good as demographic/diagnostic factors in the prediction of violence in the community, and Gray et al. (2003) report similar predictive contribution for HCR-C and HCR-R in a study of in-hospital violence. Ultimately, comparing different studies—which differ in their measures, methods, study populations and settings—is always going to be a less than ideal approach in determining whether different risk assessment strategies are required for different settings and/or populations.

On this issue, our participants’ apparent preference for using dynamic cues over static ones suggests that clinicians are willing to respond to recent calls for increased reliance on dynamic risk factors in violence risk assessment (Desmarais, Nicholls, Wilson & Brink, 2012; Douglas & Skeem, 2005; McDermott, Edens, Quanbeck, Busse & Scott, 2008). As practical constraints limited our investigation to a subset of the potential influences on risk assessment, it would be valuable to extend our investigation to

---

7Skeem and Monahan (2011) noted ‘little is known about how often or how much clinicians modify actuarial risk estimates or about the justifications they provide for such modifications’ (p. 39). Our data suggest a hypothesis in this regard: clinicians may regard actuarial tools as generic tools that require modification in light of the specific features of the kind of violence or the setting (e.g., location) that is of interest. In a similar vein, Cahill and Rakow (2012) note that risk assessments may draw on different cues depending upon whether the target event (e.g., violence) is being assessed for first occurrence (initiation), continuation, re-occurrence or escalation of the event.
examine how clinicians’ judgements would utilize the dynamic risk factors that these authors identify (e.g., impulsivity and anger). Similarly, it would be valuable to investigate clinician’s preferences for incorporating protective factors into risk assessments: first, because these may be important for management planning in routine clinical practice; and second, because these have only recently begun to be included in structured risk assessment tools (Desmarais et al., 2012). Protective factors such as strong social support, strong social bond or resilient personality traits have been shown to reduce the risk of violence (Ullrich & Coid, 2011). An assessment of risk factors as well as preventative factors should allow for a more complete and valid assessment of the risk of future violence. A further improvement to our study would be to add cues to the case scenarios that research suggests are not predictive of violence risk assessment. Our focus was on how clinicians used cues that are known to have some validity; however, it may be just as dangerous to use an invalid cue as it is to fail to use a valid cue. It would therefore be valuable for future research to determine if and when clinicians incorporate cues that do not appear on structured tools into their risk assessments—and we suggest that clinical judgement analysis is a suitable vehicle for doing so.

A further line of future research is suggested by the finding that both violence frequency and severity were seen as important for assessing danger in hospital (Table 4), whereas frequency of violence dominated assessments for the probability of in-hospital violence (Table 3). This highlights that, even when risk is assessed for a given setting, different components of risk (Dowie, 1999) or different specific events (Cahill & Rakow, 2012) may be considered by the risk assessor. It would be valuable to determine which kind of risk assessment (e.g., probability versus danger) predicts clinicians’ preferences for action and intervention more strongly. For instance, similar risk assessments by different clinicians do not always lead to similar recommendations for action (Cheyne et al., 2012) —although the extent of such discrepancies could depend upon which kind of risk assessment is made. For each of these possible lines of research, it would be important to recruit larger samples of participants than we were able to, in order to allow for better generalizability.

This, therefore, is an area ripe for further research: to determine the validity of our participants’ clinical intuitions that different settings (e.g., community or hospital) require violence assessments that differ in important ways. If these intuitions prove to be appropriate, there are several practical implications. Different risk assessment tools, or separate subscales of these tools, may have to be selected according to the setting or outcome (e.g., violence type) under consideration. Alternatively, existing tools may have to be tailored to allow for better setting-specific or outcome-specific predictions. Or, this may be an (unusual) instance where (generic) actuarial assessments may benefit from adjustment by the clinician to take account of the specific circumstances under consideration (Lawrence, Goodwin, O’Connor & Önkål, 2006).

ACKNOWLEDGEMENTS

This paper is an equal collaboration between the two authors; the order of authorship is arbitrary. We thank Rebecca Wright for assistance with preparing our materials and Nick Sevdalis and Jo Smith for comments on an earlier version of this manuscript.

REFERENCES


APPENDIX 1

Example of case scenario with the three risk-assessment questions as presented to participants

You are asked to assess a 32 years old male for his six monthly CPA* meeting. He has been a patient in this psychiatric hospital for nearly 6 months. The patient has been previously diagnosed with psychopathy. The current presentation of violent episodes has a relatively short history, going back approximately 10 months before this admission. On average, he exhibited violent episodes once every 6 months during that time. The majority of those episodes were classed as mild. The patient presented behavioral problems early at school. His first violent offence was at the age of 16. Since admission to psychiatric hospital he exhibited a violent episode once in the last 6 months. The episode was described as mild in nature.

Please present your assessment of this case by answering following questions:

1. **What is the probability that this patient will exhibit violent behaviors in the next 6 months if discharged to the community?** (Please mark your response on the line):

   Low   Medium   High

2. **What is the probability that this patient will exhibit violent behaviors in the next 6 months if he remains at his current hospital?** (Please mark your response on the line):

   Low   Medium   High

3. **Currently, what level of danger does this patient pose if he remains at his current hospital?** (Please mark your response on the line):

   Low   Medium   High

*CPA = Care Programme Approach
Box 1. Template of a narrative scenario including all levels of cues (in bold)

You are asked to assess a 31 - 35 years old male for his six monthly CPA* meeting. He has been a patient in this psychiatric hospital for nearly 6 months. Previously conducted psychopathy assessments showed that he scored low and therefore has no diagnosis of psychopathy / The patient has been previously diagnosed with psychopathy. The current presentation of violent episodes has a relatively short / long history, going back approximately / as far as 8-12 months / 4-7 years before this admission. On average, X exhibited violent episodes once every 6 months/ every month / a week during that time. The majority of those episodes were classed as mild / moderate / severe. He showed no behavioral problems at school / He presented behavioral problems early at school. His first violent offence was at the age of 21 -23 / 13 - 16. Since admission to psychiatric hospital, on average X exhibited a violent episode once in the last 6 months / every month / a week. The episodes were described as mild / moderate / severe in nature.

*CPA = Care Programme Approach