Editorial

Predictive properties of risk assessment instruments following self-harm[†]

David Owens and Rachael Kelley

Summary

This month's *BJPsych* publishes two important studies concerned with the use of risk assessment scales after self-harm, one a systematic review and the other a multicentre cohort study. We agree with the authors: that each study adds weight to the existing evidence that points towards avoiding the use of such scales in clinical practice.

David Owens (pictured) is an Associate Professor of Psychiatry in the Institute of Health Sciences at the University of Leeds. Rachael Kelley is a Research Fellow, currently at the Centre for Dementia Research, Leeds Beckett University.

Background

When it comes to the clinical care of people who have attended the general hospital because of self-harm, are routine risk assessment scales on their way out? Last year, in an editorial right here in the *BJPsych*, Mulder and colleagues wrote by no means the first obituary for risk assessment scales when they were asked, rather as we have been, to comment on a newly published paper.[1] They were weighing up a systematic review that had spun out of the meticulous compiling and analysis that the National Institute for Health and Care Excellence (NICE) does so well. That review showed that none of the various risk factors (such as gender, poor health or history of psychiatric contact), nor the three risk scales for which they could find data (the Beck Hopelessness Scale, the Suicide Intent Scale and the Scale for Suicide Ideation), are of any practical value to the assessing clinician when it comes to predicting subsequent suicide among people who have attended hospital following self-harm.[2] Chan *et al* said that 'The idea of risk assessment as risk prediction is a fallacy and should be recognised as such'.[2] The editorial writers went one step further: 'Our current preoccupation with risk prediction has the potential to harm patients, clinicians and the organisations in which they work [creating] a sense of unease among clinicians and a culture of blame when things go wrong'.[1]

Mounting evidence of the poor predictive power of risk scales

In this month's *BJPsych* are two studies that take another look at the matter: a large prospective cohort analytic study from the UK,[3] and an international collaborative systematic review with meta-analysis.[4] Again, both studies focus on psychosocial risk scales, although rather more of them – 7 in the cohort study, and 39 in the systematic review (from 53 separate primary studies). The outcomes include repeated non-fatal self-harm in both the present studies, with suicide an additional outcome in the review. The cohort study deals only with prediction following an episode of non-fatal self-harm whereas the systematic review

also examines the use of these scales among other psychiatric patients. Because we already know, as above, that the notion of there being some clinical value from this kind of risk assessment is a fallacy, it seems inevitable that the new expanded research would draw similar conclusions.

Within the two new research reports are many findings that confirm the clinical futility of trying to use scales for the prediction of suicide or repeat self-harm. For example, the psychosocial instruments show only modest or poor positive predictive values – 39% for self-harm plus suicide (only 28% and 4% respectively for self-harm and suicide, if examined separately).[4] Positive predictive value is widely regarded as the measure with the most straightforward clinical meaning – the proportion of the patients who have tested positively (scoring above a threshold) who go on to experience the outcome. Plainly, a 4% predictive value means that 96% of people identified as at increased risk would not die by suicide in the follow-up period. The predictive value for non-fatal repetition, approaching 40% when combined with suicide, looks considerably better but there is something illusory about its apparent benefits when we reflect that, regardless of any testing, around 20–30% of these patients will repeat; positive predictive value is strongly associated with the incidence of the outcome event, which is rare when the event is suicide but common when it is non-fatal repetition.

The relation between positive predictive value and outcome incidence is a further problem, acknowledged by the authors, when pooling the predictive values in the meta-analysis. Each primary prediction study in the review has its separate and differing case-mix and follow-up duration, thereby rendering the pooled predictive value one that is derived from widely differing outcome incidences. The pooling of these values in a validity meta-analysis is much more of a methodological problem than it is in clinical trial meta-analysis. In a review of trials, the main finding for each trial is a comparison between two trial arms that have a common baseline outcome incidence, although those outcomes may differ widely between one trial and another. Pooling the comparison of the two arms of each trial is not invalidated by the finding of a range of baseline outcomes. When it comes to the pooling of predictive values in a validity meta-analysis, however, we find that some studies researched subgroups of patients for whom there was a high outcome incidence, and some undertook long follow-up periods, together leading to pooled positive predictive values that are likely to be substantially higher than the performance level that would be expected in regular clinical practice.[4]

If the systematic review looks unsupportive of risk-scaling, the cohort study throws an even more unflattering light on the

[†]See pp. 387–395 and 429–436, this issue.

predictive power of risk scales in self-harm.[3] In five large English teaching hospitals, patients referred to multidisciplinary liaison psychiatry services for psychosocial assessment after self-harm were administered a structured assessment that contained the questions that make up five well-known named assessment scales – many but not all of the items were ones routinely asked but some additional questions needed to be included in the research assessment interview. They also added a clinician-rated global evaluation scale and a patient-rated version. In total, 514 patient-episodes of self-harm were assessed in this way across the five centres – each patient contributing to the validity appraisal of each of the seven scales – and each patient was followed up, using the local hospital databases, for 6 months to identify whether he or she repeated self-harm. At 30% repetition in 6 months, the positive predictive values, using established cut-off points, varied from a meagre 13% (the Modified SAD PERSONS scale) to 47% for the clinician's global scale. So, the scales' predictive values ranged from worse-than-useless to providing some modest predictive advantage, although the clinician's rating was the scale that offered the best forecast (and the patient's own global rating was the next best).

The authors went on to use a technique familiar to these kinds of evaluations: they recalculated cut-off points *post hoc*. That is, they used the findings to determine the best possible cut-off threshold for their particular study sample. It is important to emphasise that the revised validity is thereby one that maximises the validity metrics; if the study were repeated in another, independently assembled sample of patients (or in the real world) the scales would function a little or a lot worse.[5] In these ideal calculations, the self-harm prediction scales are crowded together and provide a range of positive predictive values between 33% and 47%; at the top end of this range of values they accomplish no more than we saw in the systematic review.[4] Another useful way of judging validity is the plotting of receiver operating characteristic (ROC) curves and the calculation of areas under the curves of competing scales; the seven scales here show poor-to-worthless performance. In short, no scale provides a reasonably accurate prediction of repetition of self-harm.

The way forward – needs assessment

What then is to be done? The research cited above and the earlier editorial[1–4] consistently set out a clear recommendation: risk assessment scaling should not form the basis of clinical care, and the use of detailed risk assessment scales should be curtailed lest they deliver false reassurance for clinicians and managers.[1] National (NICE) guidelines already tell us: 'Do not use risk assessment tools and scales to predict future suicide or repetition of self-harm [or] to determine who should or should not be offered treatment or who should be discharged'[6] – yet use of these tools remains widespread.[7] Instead, thorough psychosocial assessment is recommended after self-harm – to 'move away from assessment models that prioritise risks at the expense of needs'[2] and to focus on the person and the particular circumstances, characteristics and meanings that seem likely to have precipitated his or her suicidal ideation and behaviour. Needs assessment is a sufficiently apt term for such an approach and it is probably much closer to what patients actually want.

Evidence, largely qualitative, about recipients' views of the psychosocial assessment process offers some straightforward messages, although what people think specifically about being in receipt of risk-scaling is not known. There are copious reports of dissatisfaction with any sense of being processed, with stock questions in particular deemed to constitute a superficial assessment, for example: 'I could say, "how are you feeling now?" and get paid for it . . . in that sort of work you've got to have your own opinions and I think that's what makes the better ones that don't stick to the standard questions, yes and no answers;'[8] and 'it's like they're pre-programmed to ask these questions irrespective, you could go with your head dropping off and they'd still be asking these questions'.[8]

It is plain that patients want a space where it is safe to be emotionally distressed while retaining some sense of privacy.[8] When it comes to investigations, physical or psychological, people want clear explanations of why they are being done, what is involved, what the findings and implications are, and want to discuss what further symptoms they might expect as a consequence of the self-harm and of any treatment given.[8] People also want to share in decision-making about their future care, with reasonable attention paid to their personal preferences.[9,10] It is not that people who have attended hospital as a consequence of self-harm want distinct or exceptional treatment. Instead, they want the same level of clinical care that might be expected by anyone else in the emergency department or general ward, delivered with the same level of openness, warmth and respect – although accompanied by acknowledgement of their fragile emotional state. They also want assurance that they are not viewed as time-wasters or attention-seekers.[11] Care delivery, including the psychosocial assessment process, needs to be sufficiently compassionate. People are relieved to have their painful mental state taken seriously and when the nurse or doctor legitimises feelings of distress, it can be a first step in dealing with the intense negative emotions that preceded the self-harm.[12] People who have self-harmed know that there will be some routine questioning and a necessary assessment process before they can go home[12] and they know it will be required before arrangements that they hope will be helpful can be put in place: '[I'm] hugely grateful that I've got the help, it's made a whole world of difference . . . people are phoning me, keeping me informed, my care people are coming, I know that within the next couple of weeks, I will have the support I need'.[12]

Unfortunately, there is a substantial flaw in any recommendation concerned with improvements in the quality and focus of psychosocial assessments, revealed by a recent survey of hospital services for self-harm in 32 hospitals in England by Cooper and colleagues.[13] The study reveals a sad state of affairs regarding hospital services for self-harm: the median figure for the proportion of people receiving a psychosocial assessment from a mental health professional following self-harm was only 58%, and was as low as 22% at worst – worrying rates given the strong links between self-harm and subsequent suicide. This survey was a re-evaluation of the same hospitals in a similar survey almost 10 years earlier and there seemed to have been no improvement on the low levels of adequate assessment seen then, although there was some evidence of improved service structures. Further corroboration of this habitual deficiency in the services in England can be found in the present cohort study:[3] in the teaching hospitals where the research was carried out, psychosocial assessment was only received by 45% to 77% of people. It is hard to see how a focus on individual patients' needs, as recommended here and elsewhere, can be achieved when many (and in some places most) patients do not receive an assessment at all, let alone a proper evaluation of those needs.

David Owens, MD, MRCPsych, Institute of Health Sciences, University of Leeds, Leeds, UK; **Rachael Kelley**, RMN, MSc, Faculty of Health & Social Sciences, Leeds Beckett University, Leeds, UK

Correspondence: David Owens, Leeds Institute of Health Sciences, Worsley Building, Clarendon Way, Leeds LS2 9NL, UK. Email: d.w.owens@leeds.ac.uk

References

1 Mulder R, Newton-Howes G, Coid JW. The futility of risk prediction in psychiatry. *Br J Psychiatry* 2016; **209**: 271–2.

2 Chan MKY, Bhatti H, Meader N, Stockton S, Evans J, O'Connor RC, et al. Predicting suicide following self-harm: systematic review of risk factors and risk scales. *Br J Psychiatry* 2016; **209**: 277–83.

3 Quinlivan L, Cooper J, Meehan D, Longson D, Potokar J, Hulme T et al. Predictive accuracy of risk scales following self-harm: multicentre, prospective cohort study. *Br J Psychiatry* 2017; **210**: 429–36.

4 Carter G, Milner A, McGill K, Pirkis J, Kapur N, Spittal MJ. Predicting suicidal behaviours using clinical instruments: systematic review and meta-analysis of positive predictive values for risk scales. *Br J Psychiatry* 2017; **210**: 387–95.

5 Sackett DL, Haynes RB. The architecture of diagnostic research. *BMJ* 2002; **324**: 539–41.

6 National Institute of Health and Care Excellence. *Self-Harm. The NICE Guideline on Longer-term management. National Clinical Guideline Number 133.* British Psychological Society & Royal College of Psychiatrists, 2011.

7 Bolton JM, Gunnell D, Turecki G. Suicide risk assessment and intervention in people with mental illness. *BMJ* 2015; **351**: h4978.

8 Horrocks J, Hughes J, Martin C, House A, Owens D. *Patient Experiences of Hospital Care Following Self-harm – A Qualitative Study.* University of Leeds, 2005.

9 Taylor TL, Hawton K, Fortune S, Kapur N. Attitudes towards clinical services among people who self-harm: systematic review. *Br J Psychiatry* 2009; **194**: 104–10.

10 Palmer L, Blackwell H, Strevens P. *Service Users' Experiences of Emergency Services Following Self-Harm: A National Survey of 509 Patients.* CCQI, Royal College of Psychiatrists, 2007.

11 Owens C, Hansford L, Sharkey S, Ford T. Needs and fears of young people presenting at accident and emergency department following an act of self-harm: secondary analysis of qualitative data. *Br J Psychiatry* 2016; **208**: 286–91.

12 Hunter C, Chantler K, Kapur N, Cooper J. Service user perspectives on psychosocial assessment following self-harm and its impact on further help-seeking: a qualitative study. *J Affect Disord* 2013; **145**: 315–23.

13 Cooper J, Steeg S, Bennewith O, Lowe M, Gunnell D, House A, et al. Are hospital services for self-harm getting better? An observational study examining management, service provision and temporal trends in England. *BMJ Open* 2013; **3**: e003444.

---

psychiatry in literature

The Counterfeiters by André Gide

Alistair Stewart

Sexuality (homo and hetero), youth and age, seduction, corruption, religion, respectability and hypocrisy, uncertain parentage, masturbation, innocence and cynicism, adultery, mental illness, art, appearance and reality, suicide, partying and self-discovery – André Gide packed all this into *The Counterfeiters*, his novel of 1925, set among the Parisian bourgeoisie of the early 20th century. One innocent caught up in this tangled and dangerous web is a vulnerable, troubled and lovable boy of 13 called Boris. His mother is a highly strung travelling singer; his father the man with whom she eloped, but who is now dead; he is pined for by a grandfather, an ailing, unhappy man who has never seen him. Boris has been entrusted to the care of a Polish psychoanalyst, Mme Sophroniska, who is trying to release him from a state of regressed and erratic behaviour accompanied by a private language. She in turn describes her daughter Bronja, with whom Boris has formed an intense bond, as his 'real doctor'.

The character of Sophroniska is based on the Polish psychoanalyst Eugenia Sokolnicka, pupil of Jung and analysand of Freud in Vienna, and then, more successfully, of Ferenczi in Budapest. A complex personality, she became a key figure in the introduction of Freud's methods into France and a pioneer of child psychoanalysis. She was known to André Gide and treated him, although he only attended for six sessions; possibly he was only seeking material for his writings. He certainly drew for Boris's story on some superficial details of Sokolnicka's account, published in 1920, of her analysis of a 10-year-old boy with an obsessional neurosis. And although the cases of the real and the fictional child are quite different, there are clear echoes in Sophroniska's approach of Sokolnicka's careful, undogmatic and practical method, and her emphasis on the importance of basic sex education, as reflected in her paper, which incidentally is not difficult to find.

Sophroniska is described by Edouard, a somewhat ambiguous and morally complacent author, who meets her and Boris and Bronja at a resort in the Swiss Alps. She gently teases him about his artistic pretensions, and explains her own procedure as follows: 'My role is to allow things to emerge, and above all not to suggest anything. This requires extraordinary patience'. Her elucidation and formulation of Boris's troubles, her treatment of him, and what subsequently happens to Boris and Bronja, form an important strand in the novel and are central to its denouement.

The title alludes to a case of circulating forged money, but also clearly has a wider resonance. As often in life, young people suffer because of the selfish, hypocritical or thoughtless behaviour of their elders; certain more cynical and malevolent individuals induce impressionable children and adolescents to do their dirty work, while keeping their own hands clean. At the same time Gide is unsentimental about how cruel children can be. It is one thing to cure a child of a neurosis, quite another to protect them from the dangers of the outside world.