

Probability and loss: two sides of the risk assessment coin

Matthew M. Large,¹ Olav B. Nielssen²

The Psychiatrist (2011), 35, 413–418, doi: 10.1192/pb.bp.110.033910

¹University of New South Wales, Sydney; ²St Vincent's Hospital, Sydney

Correspondence to Matthew M. Large (mdbl@bigpond.com)

First received 25 Dec 2010, final revision 5 Jun 2011, accepted 2 Aug 2011

Summary Risk assessment has been widely adopted in mental health settings in the hope of preventing harms such as violence to others and suicide. However, risk assessment in its current form is mainly concerned with the probability of adverse events, and does not address the other component of risk – the extent of the resulting loss. Although assessments of the probability of future harm based on actuarial instruments are generally more accurate than the categorisations made by clinicians, actuarial instruments are of little assistance in clinical decision-making because there is no instrument that can estimate the probability of all the harms associated with mental illness, or estimate the extent of the resulting losses. The inability of instruments to distinguish between the risk of common but less serious harms and comparatively rare catastrophic events is a particular limitation of the value of risk categorisations. We should admit that our ability to assess risk is severely limited, and make clinical decisions in a similar way to those in other areas of medicine – by informed consideration of the potential consequences of treatment and non-treatment.

Declaration of interest None.

Assessment of the probability of future harm, often referred to as a 'risk assessment', has been widely adopted in mental healthcare settings in an attempt to reduce the incidence of violence and self-harm. The aim of risk assessment is to identify individuals who are at greater risk of harm and provide those patients with a higher level of treatment and supervision, thereby reducing the incidence of harm. The term 'risk assessment' is used in a variety of ways, from the opinion of an experienced clinician about dangerousness to the use of a score derived from a checklist of factors associated with a range of harmful behaviours, particularly violence to others or suicide. The ability to assess risk is regarded as an essential skill for mental health practitioners¹ and the practice guidelines issued by governments and by professional bodies suggest that we are able to predict and prevent many forms of harm.^{2–4} Assessing whether an individual is likely to harm themselves or others is part of the mental health law in most high-income countries,^{5–7} and the routine use of structured instruments to estimate the probability of future harm, often referred to as actuarial methods, are widely believed to be a way of reducing the incidence of violence^{8–12} and self-harm.^{13–15}

Criticisms of risk assessment have been made on statistical, ethical and empirical grounds. Statistical arguments note the lack of accuracy of predictions and highlight both the high rates of false-positive predictions for most forms of harm and the failure to identify many cases.^{16–20} Ethical arguments against risk assessment include the potential for the denial of care to those classified as at low

risk⁷ and the discriminatory treatment of people who have been categorised as being at high risk but do not go on to cause or experience harm.^{21,22} Another ethical problem with risk assessment is the way it devalues patients by underestimating their capacity for choice²³ and alienates them from participating in decisions about their own care.²⁴ The empirical arguments against risk assessment include the near complete absence of published evidence that the adoption of risk assessment can result in a reduction in any form of harm. The one exception to this is a cluster-randomised trial of nine psychiatric wards that examined violence for 3 months after the adoption of structured risk assessment.⁹ This study reported a significant reduction in violence in the experimental wards. However, the two groups of wards were not matched for levels of violence before the trial and after the intervention the incidence of violence in the experimental wards returned to the level of the control wards, suggesting that the results obtained were a result of regression to the mean rather than a true effect.²⁵ Moreover, the vast majority of predictions of harm were false positives. A more recent study, also of violence in psychiatric wards found that violence risk assessment was not associated with a sustained reduction in violent incidents.²⁶

In this paper we examine the concepts involved in risk assessment and the extent to which risk assessment, in particular actuarial risk assessment, can assist clinicians in the everyday task of balancing the risk of various forms of harm and the costs of interventions designed to reduce or

prevent harm. The main argument of this paper is that risk assessment, in particular actuarial risk assessment, cannot meaningfully address the basic equation that defines risk – the sum of the product of the probabilities multiplied by the losses – and as a result can only make a very limited contribution to clinical decision-making.

What is risk assessment?

The Australian Oxford Dictionary defines risk as ‘the chance or possibility of danger, loss, injury or other adverse consequences’.²⁷ Hence risk is a combination of two key concepts: (a) the chance or possibility of harm – the probability; and (b) the nature and extent of the harm or injury – the loss.

The theory of probability that underpins all risk assessment originated in correspondence between Pascal and Fermat in 1652.^{28,29} Pascal is thought to have sponsored the authors of *La Logique, ou L'art de Penser* (Logic or the Art of Thinking) published from the Port-Royal Monastery in the 1660s, which, in its final chapter ‘Belief in future contingent events’ emphasises why ‘Fear of an evil ought to be proportionate not only to the magnitude of the evil but also to the probability of occurrence’.³⁰ The concepts of probability and loss were re-stated in the early eighteenth century by Abraham de Moivre, the inventor of the bell curve and the statistical concept of variance, in his treatise on gambling, *De Mensura Sortis* (On the Measurement of Chance), in which he wrote ‘the price of a game is the loss involved multiplied by its probability’.³¹ The same definition of risk is used today in insurance, finance and engineering.³² can be found in various forms throughout the development of probability theory²⁹ and in contemporary definitions, including in Wikipedia.³³ Hence the mathematic definition of risk is: $R_i = L_i P(L_i)$, where R_i is the risk, L_i is the loss, and $P(L_i)$ is probability of the loss. De Moivre and Wikipedia also agree that when multiple independent risks are present the total risk is the sum of the individual risks and can be expressed as $R_{\text{total}} = \sum_i L_i P(L_i)$. Potential losses are rarely only financial, but the balance between premiums and payouts in insurance provides a clear illustration of the principles of risk assessment.³⁴

Risk assessment in mental health

The aim of any estimation of risk in mental health is to prevent harm to self or others, and the term ‘risk assessment’ is usually used to describe methods of estimating the probability of violence to others, self-harm and suicide. Psychiatrists have an understandable desire to protect both their patients and people who might be harmed by them, as well as a requirement to abide by local mental health laws and statutory and common law duties of care to patients and to others.³⁵ Clinical assessment remains the most common way of estimating the likelihood of future harm.³⁶

In mental health, the costs and losses involved cannot always be defined in financial terms, but the core principles of risk assessment are the same as those of insurance. For example, the risk of suicide for a patient’s family is the

numeric probability of suicide multiplied by the factors such as the grief, suffering and financial hardship should suicide occur. The risk of suicide for a patient’s clinician is the probability of the suicide multiplied by the personal, legal and professional consequences, as well as the financial cost covered by the medical defence insurer.

Clinical methods of risk assessment

In most instances mental health practitioners rely on their knowledge and experience to estimate the possibility of future harm. This amounts to an informed guess based on the individual’s presentation, history and circumstances. Although widely used, clinical methods have a subjective element, lack transparency³⁶ and have been repeatedly shown to be inferior to more systematic methods for estimating the likelihood of violence and offending.^{37,38}

Clinical risk assessment depends on the subjective interpretation of the person’s condition and situation, and is often hampered by the inability of a clinician to consider all the data available about an individual. Another problem is the natural human tendency to pay undue attention to high-loss events with low probability (such as homicide or suicide), described in the Nobel Prize winning work of Kahneman & Tversky.³⁹ An example in mental health is the fear and community reaction to the homicide of strangers by people with psychosis. Although the incidence of stranger homicide by people with schizophrenia is about 1 in 15 million people per year – making them one of the rarest causes of death – these events attract huge media coverage and have resulted in changes to both clinical practice and to mental health law in several jurisdictions.⁴⁰

The heuristic nature of clinical risk assessment, in which the steps to reaching a decision are not always clear even to the clinician, means that the basis for making decisions are difficult to explain and defend after adverse events, especially at formal inquiries conducted with the bias of hindsight.⁴¹

Actuarial methods and structured clinical judgement

Actuarial risk assessment assesses the probability of an adverse event by scoring patient characteristics according to the presence or absence of a predetermined set of risk factors. Actuarial methods use scales derived from factors found to be associated with adverse events in previous research to categorise patients into high- or low-risk groups. For example, a hypothetical five-point scale to assess the probability of future violence might comprise items for: (a) male gender; (b) young age; (c) past violence; (d) substance use; and (e) the presence of psychotic illness. A patient rated to have four or five of these features might be categorised as being at higher risk than people with a score of three or less.

As well as being more accurate than clinical methods, actuarial methods have a higher interrater reliability, are more transparent and require less expertise. However, actuarial instruments do not perform as well in subsequent trials when compared with the original sample because of the high likelihood of chance findings being included in the

initial model, and the inevitable differences between the sample from which the model was derived and subsequent groups of patients.³⁸ Moreover, most of the established risk factors, such as gender and a history of offending, cannot be modified and do not provide a good guide to clinical decisions such as whether a patient who has committed an act of violence prior to admission can be safely discharged from hospital.

Actuarial risk assessments produce a score, almost always in the form of an interval scale generated by adding a score from each item of risk found to be present. In practice, a cut-off score must be chosen if the risk assessment is to be used to guide clinical decisions, placing the patient being assessed into a category deemed at high or low risk. Risk assessment conducted in this way might inform decisions to admit or discharge a person, prescribe electroconvulsive therapy, commence clozapine or administer antipsychotic medication by long-acting injection.

The hope that clinical methods, that are tailored to individual circumstances, can be effectively combined with more reliable actuarial methods has led to the development of structured clinical judgement that combines features of both actuarial and clinical judgement regarding the patient's mental state and immediate plans.³⁸ However, regardless of the method used to assess risk, the clinician still has to choose a threshold at which to intervene.

Hence, all forms of risk assessment divide groups of patients into higher- and lower-risk categories. A binary risk estimation result in a 2 × 2 contingency table of high and low probability categorisations and harm would or would not occur. A 2 × 2 contingency table forms the basis of the following discussion – but similar arguments can be made when patients are categorised into low-, medium- or high-risk groups or even on the basis of each increment of a risk scale.

The probability arm of risk assessment

The contingency table of high- and low-risk categorisation and the actual occurrence of future harm produces four outcomes:

- (1) true positive (TP), where a person categorised as being at high probability of harm will commit that harm;
- (2) false positive (FP), where a person categorised as being at high probability of harm would not commit the harm;
- (3) true negative (TN), where a person categorised as being at low probability of harm would not commit the harm; and
- (4) false negative (FN), where a person categorised as being at low probability of harm will commit that harm.

These four outcomes can generate the well-known statistics associated with risk assessment. Sensitivity ($TP/(TP + FN)$) is the proportion of correctly identified cases detected by the instrument, and specificity ($TN/(TN + FP)$) is the proportion of correctly identified non-cases. There is always a trade-off between sensitivity and specificity.

In other fields of medicine, sensitive tests (such as a chest x-ray) are useful for screening and can be followed up with specific or diagnostic tests (such as a biopsy). However, there are no diagnostic tests for future behaviour and we

have to be satisfied with a range of sensitivity and specificity combinations based on different cut-off scores. For example, in our five-point violence scale, a specific but insensitive test of future violence might require the patient to have a score of five, whereas a score of three or more would be more sensitive but less specific.

When the range of sensitivities is plotted on the y-axis of a graph against the corresponding specificities on the x-axis, a curve known as the receiver operator curve (ROC) is formed. The term comes from a calculation of the ability of radar (receiver) operators to detect incoming enemy aircraft after the 1941 attack on Pearl Harbour.⁴² A related statistic is the area under the (receiver operator) curve (AUC). The AUC is a number between 0 and 1 and is the probability that a randomly selected patient who goes on to commit future harm will have had a higher score than a randomly selected patient who does not commit harm. The AUC has become a common measure of the accuracy of a risk assessment.¹¹ However, a particular sensitivity and specificity on the ROC curve must be chosen if they are to guide treatment decisions. The sensitivity, specificity, ROC and AUC are necessary to assessing the psychometric properties of actuarial instruments, but they are not sufficient in themselves, because the performance of a risk assessment instrument also relies on the base rate of the adverse event.

The accuracy of high-risk categorisations is measured by the positive predictive value (PPV). Bayes' theorem tells us that the probability of harm in high- and low-risk populations is measured by positive predictive value ($PPV = TP/(TP + FP)$) and negative predictive value ($NPV = TN/(TN + FN)$). Both PPV and NPV are highly dependent on the base rate of the event being predicted. For example, an outstanding risk-assessment instrument (sensitivity 0.8 and specificity 0.8) will have a PPV of 0.3 for a common event such as a physical assault with a base rate of about 1 in 10 per annum, but would have a PPV of 0.0004 for homicide committed by a person with schizophrenia, which has a base rate of 1 in 10 000 per annum.^{18,43} Even if the (improbably high) sensitivity and specificity were applied to the prediction of both assault and homicide, a third of the predictions of assault would prove to be correct, but as few as 1 in 2500 predictions of homicide would eventuate.

The PPV is the crucial measure of the statistical accuracy of an instrument when it is applied to a specific population. Positive predictive value is a more useful statistic for the estimation of risk than sensitivity, specificity or AUC, because P (probability) in the risk estimation equation is the same as PPV such that in practice $Risk = PPV \times Loss$. It is then clear that the value of P is not the same as a risk assessment because the losses must also be known. In fact, although risk is often used synonymously with probability, the units of risk are the units of the loss – whether it is in terms of injury, loss of life, financial or damage to professional reputations. Additional treatments, closer supervision and more restrictive care would be hard to justify on the basis of the results of a risk assessment if the loss involved amounted to the effect of verbal aggression. It would also be hard to justify the detention

of 35 000 people with schizophrenia for a year in order to possibly prevent one homicide of a stranger.^{18,40}

The loss arm of risk assessment

The problem of the range of harms

The mathematical superiority of actuarial over clinical methods for predicting particular harms would appear to support the use of risk-assessment instruments, even if the final decision is a clinical judgement. However, the question then arises: which instrument should be used? Violence and self-harm are quite separate potential losses, with very different risk factors. Take the example of a person admitted with first-episode psychosis, who, like many other patients, might be considered to be at increased risk of both violence to others and suicide. Two recent systematic reviews with meta-analysis cast some light on the risks faced by this individual. In the first review of risk factors for violence to others in first-episode psychosis we found that male gender, younger age and substance use were risk factors for violence.⁴⁴ In the second review of risk factors for in-patient suicide, age and gender were not associated with suicide and there was a trend for a history of substance use to be protective.⁴⁵ Depressed mood was found to be protective against violence but was not surprisingly a risk factor for suicide. Therefore, it is likely that a patient with an increased probability of self-harm relative to others also has a lower probability of violence to others and vice versa. A further complication is that harmful acts of differing severity such as minor acts of aggression, acts of physical violence and violence causing injury to others have different associations and hence require different risk-assessment instruments.^{44,46,47} A good example is the presence of mania, which is associated with aggression and minor violence, but is not strongly associated with severe violence or homicide.^{48,49}

Further research might produce an instrument that could assess a variety of harms of differing levels of severity and different base rates. However, at present there is no actuarial risk assessment method that is able to assess the possibility of more than one type of harm, adjust the predictive validity in line with the base rate of each harm, or consider the various risk factors associated with the varying levels of self-harm and harm to others.¹⁸ The problem of cumulative multiple risks associated with different types of harm, expressed in the equation $R_{\text{total}} = \sum_i L_i p(L_i)$ has not been addressed in any actuarial method of risk assessment.

The problem of the extent of losses

Although risk-assessment instruments can make a distinction between those at higher and lower probability of harm, they cannot necessarily make a meaningful estimation of the extent of the actual loss. Some studies of violence such as the MacArthur study do make a distinction between serious and less serious forms of violence,⁵⁰ but none of the widely used tools for the prediction of violence, self-harm and criminal offending makes an attempt to quantify the losses arising from violence of differing severities.

There are a number of studies examining the risk factors associated with the comparatively rare but well-defined loss of a completed suicide.⁵¹ However, most

consider a combination of different types of episodes of self-harm or even threatened self-harm,¹³ or a range of violent acts of varying severity.⁵² There is great variation in both the intention and outcome of violent actions, from a minor shove to an actual homicide. Moreover, similar harms can have very different costs, for example consider that two people taking overdoses of the same drug at the same dose can have quite different medical outcomes, not everyone who is assaulted develops post-traumatic stress disorder and an apparently minor assault on a vulnerable older person can result in death.

Finally, there will always be a subjective element to the quantification of loss, in part because of individual differences in value systems. The estimation of a loss by a patient might differ from the loss estimated by the clinician, an in-patient suicide might be judged to be a less serious loss than an in-patient homicide, although both involve the death of a patient, and the rare homicide of a stranger might be viewed to be of greater concern than the more common homicide of a family member.⁴⁰

Risk assessment is popular with politicians and service managers, because it appears to present a solution to highly publicised catastrophic events, such as the tragic homicide of Jonathan Zito by a stranger with schizophrenia, Christopher Clunis.⁵³ Hence it is ironic that the widespread adoption of risk assessment has been in response to very rare high-loss events that cannot be predicted.

Discussion

Previous criticisms of risk assessment have noted the statistical limitations, ethical problems and the absence of empirical evidence that risk assessment has ever been shown to reduce harm. This examination of the components of risk raises two further major problems with risk assessment – the inability of actuarial instruments to consider the range of harms associated with mental illness and therefore the total risk, and the inability to assess the extent of loss arising from an adverse event.

Hence the term risk assessment is really quite misleading when used in the context of mental healthcare, because risk assessment in its current form does not estimate the sum of the potential harms or the extent of potential losses, according to the ordinary definition of risk. If we are unable to estimate the range of outcomes, how can we hope to balance the risks with appropriate measures to prevent harm – the psychiatric equivalent of an insurance premium?³⁴ This premium, in the form of interventions such as increased medication, closer supervision and longer detention in hospital, is paid by all those patients categorised as high risk and by the health services providing those treatments, on the assumption that the costs of false-positive assessments are justified for the wider good.

Like other doctors, psychiatrists have to assist patients and their families to strike a balance between the benefits and harms associated with medical treatment. This includes the harms of non-treatment. When a cardiologist elicits a history of cardiac symptoms and risk factors for heart disease, a discussion usually takes place between the doctor and patient about further investigations, treatment and ways of preventing disease. There is an element of subjectivity in the estimation of that risk, because two

people with similar illnesses and mental capacity for decision-making can make very different choices about how to act on the advice they receive. In mental health settings, risk assessment should inform a similar discussion with any patient who is competent to consider advice. However, it is rarely used in this way,²⁴ even though the patient's own judgement about their propensity for violence has been shown to have a similar level of accuracy to actuarial risk assessment.⁵⁴

It is unlikely that clinical methods can satisfactorily achieve the goal of an accurate and comprehensive prediction of all forms of harm either, not because we are unable to consider a range of harms and losses, but because humans are poor at making choices based on information about risk, particularly when considering rare but highly adverse outcomes.³⁹ However, the clinician is in a position to advise the patient about the range of adverse events as well as the limitations of prediction, and to discuss these with the patient's family and with colleagues. In this way, the doctor can attempt to balance the risks associated with the patient's condition with the costs and benefits of psychiatric treatment.

Structured clinical judgement has been proposed as a solution to the limitations of both actuarial and clinical risk assessment. Structured clinical judgement allows the assessor some latitude in their judgement of the weight to be placed on the score derived from actuarial instruments. However, structured clinical judgement is unlikely to be a solution to the shortcomings of either clinical or actuarial risk assessment, even if it ensures that all known factors are considered along with the patient's current condition and circumstances, because it is equally likely to reintroduce the biases and lack of transparency that actuarial methods seek to avoid, and the superiority of structured clinical judgement over other types of risk assessment has not been established.

We suggest that any type of risk assessment that categorises patients into high- and low-risk groups should not be included in clinical decision-making. This is not to say that treatment should not be offered to people who have modifiable risk factors, such as treatment for psychosis and substance use, as well as interventions to improve their social circumstances, but the decision to offer these treatments should not be based on the false assumption that an estimate of the probability of a single class of events represents the risk faced by the patient. Instead, clinical decisions should be made in the same way as in other fields of medicine, after an informed discussion with a competent patient or their proxy decision-maker.

Similarly, the capacity to give informed consent to treatment, rather than the perceived risk of harm, should be the cornerstone of non-consensual treatment. Mental capacity to consent to treatment can be operationally defined and has a high interrater agreement.^{7,55–57} Mental incapacity is the standard by which non-consensual treatment is delivered in other areas of medicine, such as to children, individuals in a coma and to people with dementia. This is not to say that the potential for violence and self-harm can be ignored, rather, that we should concentrate on providing treatment that improves the individual's decision-making ability, for example in the

form of a joint crisis plan.⁵⁸ Some patients do exhibit severe, continuous or regular violence and self-harm, and require containment in hospital for the protection of themselves and others. The clinical and ethical dilemmas posed by these individuals will not be solved by performing actuarial risk assessments, because in most cases risk assessment is redundant when the harm is ongoing. In these cases the issue is not one of prediction, and applying the results of risk assessment instruments often has the effect of preventing the clinician from considering a less restrictive form of treatment.

It has been suggested that the reluctance of mental health professionals to employ actuarial methods in clinical practice stems from the challenge they pose to the clinician's status as an expert.²⁴ Our analysis suggests that the real reason is likely to be the inability of actuarial methods to make a meaningful estimation of both the probability of multiple harms and the resulting potential losses. Although it might be too early to call off attempts to estimate risk in mental health, we should acknowledge the severe limitation in our ability to predict future harmful events to our patients, their families, medical administrators, governments and, most of all, to ourselves.

Acknowledgement

We thank Peter Arnold for his assistance with the manuscript.

About the authors

Matthew M. Large (BSc (Med), MBBS, FRANZCP) is a clinical senior lecturer at the School of Psychiatry, University of New South Wales, Sydney, Australia. **Olav B. Nielssen** (MBBS, M Crim, PhD, FRANZCP) is a psychiatrist in the Clinical Research Unit for Anxiety and Depression, St Vincent's Hospital, Sydney, Australia.

References

- 1 Weisman RL, Lamberti JS. Violence prevention and safety training for case management services. *Community Ment Health J* 2002; **38**: 339–48.
- 2 New South Wales Health. *Framework for Suicide Risk Assessment and Management for NSW Health Staff*. New South Wales Health, 2005 (http://www.health.nsw.gov.au/pubs/2005/suicide_risk.html).
- 3 Royal College of Psychiatrists Special Working Party on Clinical Assessment and Management of Risk. *Assessment and Clinical Management of Risk of Harm to Other People* (Council Report CR53). Royal College of Psychiatrists, 1996.
- 4 New Zealand Health. *The Assessment and Management of People at Risk of Suicide*. New Zealand Guidelines Group and Ministry of Health, 2003.
- 5 Anfang SA, Appelbaum PS. Civil commitment – the American experience. *Isr J Psychiatry Relat Sci* 2006; **43**: 209–18.
- 6 Dressing H, Salize HJ. Compulsory admission of mentally ill patients in European Union Member States. *Soc Psychiatry Psychiatr Epidemiol* 2004; **39**: 797–803.
- 7 Large MM, Nielssen O, Ryan CJ, Hayes R. Mental health laws that require dangerousness for involuntary admission may delay the initial treatment of schizophrenia. *Soc Psychiatry Psychiatr Epidemiol* 2008; **43**: 251–6.
- 8 Roaldset JO, Hartvig P, Bjorkly S. V-RISK-10: validation of a screen for risk of violence after discharge from acute psychiatry. *Eur Psychiatry* 2011; **26**: 85–91.
- 9 Abderhalden C, Needham I, Dassen T, Halfens R, Haug HJ, Fischer JE. Structured risk assessment and violence in acute psychiatric wards: randomised controlled trial. *Br J Psychiatry* 2008; **193**: 44–50.

- 10 Steadman HJ, Silver E, Monahan J, Appelbaum PS, Robbins PC, Mulvey EP, et al. A classification tree approach to the development of actuarial violence risk assessment tools. *Law Hum Behav* 2000; **24**: 83–100.
- 11 Barry-Walsh J, Daffern M, Duncan S, Ogloff J. The prediction of imminent aggression in patients with mental illness and/or intellectual disability using the Dynamic Appraisal of Situational Aggression instrument. *Australas Psychiatry* 2009; **17**: 493–6.
- 12 Carroll A. Risk assessment and management in practice: the forensic risk assessment and management exercise. *Australas Psychiatry* 2008; **16**: 412–7.
- 13 Hendin H, Al Jurdi RK, Houck PR, Hughes S, Turner JB. Role of intense affects in predicting short-term risk for suicidal behavior: a prospective study. *J Nerv Ment Dis* 2010; **198**: 220–5.
- 14 Cooper J, Kapur N, Mackway-Jones K. A comparison between clinicians' assessment and the Manchester Self-Harm Rule: a cohort study. *Emerg Med J* 2007; **24**: 720–1.
- 15 Hockberger RS, Rothstein RJ. Assessment of suicide potential by nonpsychiatrists using the SAD PERSONS score. *J Emerg Med* 1988; **6**: 99–107.
- 16 Mossman D. The imperfection of protection through detection and intervention. Lessons from three decades of research on the psychiatric assessment of violence risk. *J Leg Med* 2009; **30**: 109–40.
- 17 Nilsson T, Munthe C, Gustavson C, Forsman A, Anckarsater H. The precarious practice of forensic psychiatric risk assessments. *Int J Law Psychiatry* 2009; **32**: 400–7.
- 18 Large MM, Ryan CJ, Singh SP, Paton MB, Nielsen OB. The predictive value of risk categorisation in schizophrenia. *Harv Rev Psychiatry* 2011; **19**: 25–33.
- 19 Szmukler G. Violence risk prediction in practice. *Br J Psychiatry* 2001; **178**: 84–5.
- 20 Szmukler G. Risk assessment: 'numbers' and 'values'. *Psychiatr Bull* 2003; **27**: 205–7.
- 21 Ryan CJ. One flu over the cuckoo's nest: comparing legislated coercive treatment for mental illness with that for other illness. *J Bioeth Inq* 2011; **8**: 87–93.
- 22 Large MM, Ryan CJ, Nielsen OB, Hayes RA. The danger of dangerousness: why we must remove the dangerousness criterion from our mental health acts. *J Med Ethics* 2008; **34**: 877–81.
- 23 Mossman D. Critique of Pure Risk Assessment or, Kant Meets Tarasoff. *University of Cincinnati Law Rev* 2006; **75**: 523–609.
- 24 Langan J. Challenging assumptions about risk factors and the role of screening for violence risk in the field of mental health. *Health Risk Soc* 2010; **12**: 85–100.
- 25 Vickers AJ, Altman DG. Statistics notes: analysing controlled trials with baseline and follow up measurements. *BMJ* 2001; **323**: 1123–4.
- 26 Kling RN, Yassi A, Smailes E, Lovato CY, Koehoorn M. Evaluation of a violence risk assessment system (the alert system) for reducing violence in an acute hospital: a before and after study. *Int J Nurs Stud* 2011; **48**: 534–9.
- 27 Moore B (ed). *The Australian Pocket Oxford Dictionary, Fifth Edition*. Oxford University Press, 2003.
- 28 Hald A, de Moivre A. 'De Mensura Sortis' or 'On the Measurement of Chance'. *Int Stat Rev* 1984; **52**: 229–62.
- 29 Hacking I. *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference*. Cambridge University Press, 1975.
- 30 Arnauld A, Nicole P. *The Art of Thinking: Port-Royal Logic* (trans J Dickoff & P James; original work published 1662). Bobbs-Merrill, 1964.
- 31 de Moivre A. On the measurement of chance, or, on the probability of events in games depending on fortuitous chance (trans B McClintock; original work published 1711). *Int Stat Rev* 1984; **52**: 229–62.
- 32 Bernstein PL. *Against the Gods: The Remarkable Story of Risk*. John Wiley and Sons, 1996.
- 33 Anon. *Risk Assessment*. Wikipedia, 2010 (http://en.wikipedia.org/wiki/Risk_assessment).
- 34 Ryan CJ, Nielsen O, Paton MB, Large M. Clinical decisions in psychiatry should not be based on risk-assessment. *Australas Psychiatry* 2010; **18**: 398–403.
- 35 Anfang SA, Appelbaum PS. Twenty years after Tarasoff: reviewing the duty to protect. *Harv Rev Psychiatry* 1996; **4**: 67–76.
- 36 Steadman HJ. Predicting dangerousness among the mentally ill: art, magic and science. *Int J Law Psychiatry* 1983; **6**: 381–90.
- 37 Meehl PE. Causes and effects of my disturbing little book. *J Pers Assess* 1986; **50**: 370–5.
- 38 Buchanan A. Risk of violence by psychiatric patients: beyond the "actuarial versus clinical" assessment debate. *Psychiatr Serv* 2008; **59**: 184–90.
- 39 Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica* 1979; **47**: 263–92.
- 40 Nielsen O, Bourget D, Laajasalo T, Liem M, Labelle A, Häkkinen-Nyholm H, et al. Homicide of strangers by people with a psychotic illness. *Schizophr Bull* 2011; **37**: 572–9.
- 41 LeBourgeois 3rd HW, Pinals DA, Williams V, Appelbaum PS. Hindsight bias among psychiatrists. *J Am Acad Psychiatry Law* 2007; **35**: 67–73.
- 42 Green DM, Swets JM. *Signal Detection Theory and the Psychophysics*. John Wiley and Sons, 1966.
- 43 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.
- 44 Large M, Nielsen O. Violence in first-episode psychosis: a systematic review and meta-analysis. *Schizophr Res* 2010; **125**: 209–20.
- 45 Large M, Smith G, Sharma S, Nielsen O, Singh S. Systematic review and meta-analysis of the clinical factors associated with the suicide of psychiatric in-patients. *Acta Psychiatr Scand* 2011; **124**: 18–29.
- 46 Harris AW, Large MM, Redoblado-Hodge A, Nielsen O, Anderson J, Brennan J. Clinical and cognitive associations with aggression in the first episode of psychosis. *Aust N Z J Psychiatry* 2010; **44**: 85–93.
- 47 Milton J, Amin S, Singh SP, Harrison G, Jones P, Croudace T, et al. Aggressive incidents in first-episode psychosis. *Br J Psychiatry* 2001; **178**: 433–40.
- 48 Fazel S, Lichtenstein P, Grann M, Goodwin GM, Långström N. Bipolar disorder and violent crime: new evidence from population-based longitudinal studies and systematic review. *Arch Gen Psychiatry* 2010; **67**: 931–8.
- 49 Nielsen O, Yee N, Millard M, Large M. Comparison of first-episode and previously treated persons with psychosis found NGMI for a violent offense. *Psychiatr Serv* 2011; **62**: 759–64.
- 50 Monahan J, Steadman H, Silver E. *The MacArthur Study of Mental Disorder and Violence*. Oxford University Press, 2001.
- 51 Pokorny AD. Prediction of suicide in psychiatric patients. Report of a prospective study. *Arch Gen Psychiatry* 1983; **40**: 249–57.
- 52 Snowden RJ, Gray NS, Taylor J, Fitzgerald S. Assessing risk of future violence among forensic psychiatric inpatients with the Classification of Violence Risk (COVR). *Psychiatr Serv* 2009; **60**: 1522–6.
- 53 Ritchie J. *The Report of the Inquiry into the Care and Treatment of Christopher Clunis*. HMSO, 1994.
- 54 Roaldset JO, Bjorkly S. Patients' own statements of their future risk for violent and self-harm behaviour: a prospective inpatient and post-discharge follow-up study in an acute psychiatric unit. *Psychiatry Res* 2010; **178**: 153–9.
- 55 Dawson J, Szmukler G. Fusion of mental health and incapacity legislation. *Br J Psychiatry* 2006; **188**: 504–9.
- 56 Owen GS, David AS, Hayward P, Richardson G, Szmukler G, Hotopf M. Retrospective views of psychiatric in-patients regaining mental capacity. *Br J Psychiatry* 2009; **195**: 403–7.
- 57 Okai D, Owen G, McGuire H, Singh S, Churchill R, Hotopf M. Mental capacity in psychiatric patients. Systematic review. *Br J Psychiatry* 2007; **191**: 291–7.
- 58 Henderson C, Flood C, Leese M, Thornicroft G, Sutherby K, Szmukler G. Effect of joint crisis plans on use of compulsory treatment in psychiatry: single blind randomised controlled trial. *BMJ* 2004; **329**: 136.