Assessing the Probability of Patients Reoffending After Discharge from Low to Medium Secure Forensic Mental Health Services: An Inductive Prevention Paradox

Bob Heyman, Paul M. Godin, Lisa Reynolds and Jacqueline P. Davies

Abstract

Citizens of developed societies are troubled by those who commit ‘irrational’ crimes against the person. Reoffending after release from secure mental health services triggers particularly intense angst when amplified by media and political scrutiny. However, forensic mental health service providers making discharge decisions are required to predict the probability of patients reoffending after discharge by observing behaviour in secure institutional settings designed specifically to prevent such transgressions from occurring. This paper will explore this ‘inductive prevention paradox’ by drawing on data drawn from qualitative studies undertaken in two medium to low secure UK institutions, one providing forensic mental health and the other forensic learning disability services. The views of 56 staff members and 21 patients about risk management in forensic services were explored, and an additional 25 staff interviews were undertaken for case studies of these patients. Data analysis focussed on patient and staff responses to the inductive prevention paradox with respect to a highly sensitive issue. The wider applicability of this framework will be considered in the Discussion. It will be argued that the prognostic limitations arising from the inductive prevention paradox have been underestimated by policy-makers and in official inquiries; and that the prevailing personal risk assessment framework needs to be complemented by greater attention to the environments which patients will be discharged into.
INTRODUCTION

Mulvey had put on a penitent face and given a series of small humble nods. He knew the Governor and the Visiting Committee were watching from the gallery and he wanted to make an enduring impression. (O’Connor, 2006, p. 196)

Mulvey, the fictional nineteenth century prisoner mentioned in the opening citation, establishes trustee status in order to engineer his escape from a harsh 19th century Benthamite prison, brutally murdering a guard in the process. The present paper will explore an analogous issue, that of assessing the risk of serious reoffending by patients released from secure forensic mental health services. It will be argued that staff charged with managing this risk face the ‘inductive prevention paradox’ (Heyman et al., 2010, p. 103-104). The paradox arises whenever measures designed to reduce or eliminate a risk are implemented. An unwanted but inescapable consequence of adopting preventative measures is that risk reduction obscures direct observation of what might have happened if prophylaxis had not been attempted. In relation to the topic of the present paper, the probability of an in-patient causing harm to members of the public after discharge from a secure institution has to be estimated from observations made in an environment designed precisely to prevent such events from occurring. Put simply, because patients who are locked up cannot assault members of the public, the probability of them doing so if discharged becomes difficult to assess. The question of whether a detained forensic mental health patient would reoffend if released cannot be definitively answered through direct observation except by accepting this risk.

The wider implications of the inductive prevention paradox will be touched on in the Discussion. The data analysis presented below will be concerned with risk assessment in relation to the possible discharge of offenders from secure mental health services, an important practice and policy issue in itself, but also an illustrative example of the inductive prevention paradox. The analysis will focus on staff and patient recognition, if any, and responses to this issue. The remainder of the Introduction will address three relevant background issues: firstly, the societal attitudes which render reoffending by discharged offenders from mental health services as unacceptable risk; secondly, the UK history of medium to low secure forensic mental health services as a response to the inductive prevention paradox; and, thirdly, attempts to ‘beat the trap’ arising from the inductive prevention paradox through the use of psychometric actuarial methods. It will be argued that neither moving closed institutions closer to ‘the community’ nor resorting to psychometrics offers a plausible route to accurately quantifying the probabilities of discharged patients reoffending. However, staff who have to take decisions despite such epistemological limitations face public ignominy if a recently discharged patient reoffends. This analysis sets the scene...
for considering how staff and patients view and attempt to manage this conundrum.

**Societal attitudes to the risk of reoffending by forensic mental health patients**

Risk management for individuals who commit serious offences against the person attracts strong media and political attention in the UK and elsewhere despite the rarity of such crimes. Concern becomes especially heated when the offence is perceived as ‘dirty’, as in the case of sexual assault, particularly on children, or as driven by an ‘irrational’ motive, i.e. a belief system which the prevailing culture views as delusory. In contrast, more frequent causes of death and injury, for example those caused by traffic accidents, and even ‘rational’ offences against the person such as assaults inflicted during bank robberies, attract much less societal angst. The release of offenders who have committed irrational or dirty crimes and subsequently reoffend exponentially intensifies the already powerful threat to the social order arising from the identification of irrational and therefore unpredictable assaults. The perpetrator has not only exposed the dangerously erratic nature of human behaviour. In addition, society has registered such disturbing transgressions, but has failed to prevent them from reoccurring. The present paper is not primarily concerned with explaining societal attitudes to ‘irrational’ or ‘dirty’ offences per se, but rather with the implications of their intense unacceptability for risk assessment and management. The separation of the mentally disordered offender from modern society is now accounted for in terms of a utilitarian risk management framework even though powerful resonances of cultural pollution remain. At the same time, the difficulty of predicting individual behaviour make avoidance of reoffending difficult or impossible to achieve.

**A brief history of UK medium to low secure forensic mental health services**

UK forensic mental health services with lower levels of security than that provided in the established high security hospitals were originally conceived of with just the issue of realistic risk assessment in mind. In the terminology developed above, the main aim underpinning the design of these new services was tackling the inductive prevention paradox. The idea originated in the work of The Butler Committee (Home Office and Department of Health and Social Security, 1975). Their report reviewed the discharge of mentally disordered offenders in the aftermath of the notorious case of Graham Young who had been confined to a high security special hospital after poisoning members of his family. Young was released despite having overtly developed his poisoning skills during his hospital stay, borrowing many library books on this topic. He subsequently poisoned a number of work colleagues, causing media uproar. Butler identified systematic failure to recognise patients’ propensities to reoffend in the remote, esoteric environments of special hospitals.
At the time when the Butler Report was published, in 1975, risk assessment was a new and relatively unused procedure (Tidmarsh, 1992). In terms of the risk social science which emerged subsequently, it can be argued that the organisation of high secure institutions could generate selective perception to the point of risk blindness, with tragic but also blackly comic consequences, epitomised by the Graham Young case. Despite the report’s recommendations being widely publicised, only a few institutions providing medium to low secure accommodation were opened in the 1970s and 1980s, perhaps because state funding was not provided. After another national UK report offering a similar analysis had been published (Reed, 1992), such intermediate security facilities began to be established on a large-scale. However, the number of mentally disordered offenders detained in UK high security institutions has declined only moderately, remaining at over 1000 since the early nineties, whilst the number confined in medium to low security hospitals has increased from about 1000 in 1993 to over 2000 in 2003 (Ly and Howard, 2004). Reed presented the benefits of the latter in terms of them providing a half-way house between the high-security hospital and the community where patients could be observed and assessed in a more appropriate setting without putting the public at risk. Hence, one of the aims behind their expansion in the 1990s was to provide a more revealing environment for risk assessment in relation to the communities to which patients would eventually return. The finding that nurses’ negative stereotyping and therapeutic pessimism about patients is associated with security level (Mason et al., 2010) suggests that they can offer a more hopeful environment for patient rehabilitation. However, as will be documented in the Data Analysis section, staff charged with critical decision-making about discharge may not share the view that the probability of a patient reoffending can be validly assessed within the confines of medium and low security units.

Cases of serious reoffending by patients recently discharged from medium to low security institutions have inevitably occurred. A number of UK formal inquiries conducted retrospectively have explained such reoffending in terms of culpable service failures of communication and risk assessment. For instance, an inquiry into a killing perpetrated by a patient, John Barrett, one day after release from a medium secure unit (South West London Strategic Health Authority, 2006) criticised forensic mental health service providers for becoming too sympathetic to risky patients. As illustrated below, patients and staff who participated in the present research made a similar point when they argued that compliant patients could progress quickly through the system to discharge simply because they did not cause immediate problems. Nevertheless, careful analysis of the information available to service providers challenges the assumption underpinning the 2006 inquiry report that patient reoffending can be unproblematically predicted providing that staff try hard enough.

It might be suggested that this tragedy would not have occurred if it had not been for a single decision, to allow John Barrett out on leave from the medium secure psychiatric unit to which he had been readmitted on the day...
before he killed Denis Finnegan … Too much confidence was placed in clinical judgements unsupported by evidence and rigorous analysis. (South West London Strategic Health Authority, 2006, p.9)

Once it is assumed that service providers possess the capacity to accurately predict whether individuals will reoffend, it follows that their apparent failure to do so must arise from individual or organisational inadequacies deserving of censure. This interpretive framework thus transforms probability assessment into a moral issue. However, as well as discounting the inductive prevention paradox, such a perspective ignores the key limitation of probabilistic reasoning which, even if it can be based on induction from an adequate set of observations, is predicated on the assumption that the occurrence of adverse events is at least possible. The mindset documented above elides the (attempted) accurate calibration of probabilities with perfect prognostication. Staff thereby face a double shortfall with respect to meeting societal expectations based on the implicit assumptions that the chance of an individual reoffending can be measured, and that this metric somehow enables perfect prediction.

Confined to an institution where their riskiness is continually assessed, patients may react to being placed under constant observation. They may seek to please their clinicians in the hope of gaining more autonomy by reducing their assessed riskiness. Conversely, patients may ‘fake bad’, attempting to increase their official riskiness in order to invoke a therapeutic response, maintain self-esteem or postpone feared discharge, a phenomenon which insiders sometimes depict as ‘gate fever’ (Godin et al., 2006, p. 87). In turn, clinicians may attempt to deceive in the hope of seeing through the camouflage of self-presentation so as to uncover a presumed psychological reality, revelation of which will allow the risk of reoffending to be accurately assessed.

Service providers and users who live with this situation are likely to appreciate the problems arising from the inductive prevention paradox, an awareness documented in the Data Analysis section. The following sub-section offers an analysis of attempts to predict an individual’s probability of reoffending by means of formal risk assessment ‘instruments’. Their use is mostly complemented by multidisciplinary qualitative reviews of individual cases (Shaw et al., 2007). Nevertheless, critical analysis of the epistemology underpinning psychometric risk assessment can offer insights into the limitations of probabilistic inference in this context, consideration of which can be obscured by pragmatic imperatives and routinisation.

Risk assessment ‘tools’

The process of discharge decision-making for forensic mental health service-users is organised around case conferences and the use of risk assessment inventories, often called ‘tools’ by health service insiders, although recommendations must, in most cases, be externally ratified. The routinisation of
tool usage, a core feature of UK National Health Service culture, tends to conceal
their shaky epistemological foundations (Webb, 2012). Those used in forensic
mental health services contain sets of items on which patients are rated,
generating summative scores designed to indicate a patient’s probability of
reoffending after discharge. Analysis of their limitations sheds light on the
predictive difficulties arising from the inductive prevention paradox.

Such tools can be divided into three types (Gray et al., 2004). Actuarial measures
are based on easily encoded biographical information covering offending history
and demographic factors inductively associated with the probability of reoffending
(or, more accurately, the probability of being caught and convicted for
reoffending). An example is the Offender Group Reconviction Scale [OGRS]
by the Psychopathy Checklist, Screening Version [PCL-SV] (Hart, Cox and Hare,
1995), focuses on mental states presumed to give rise to offending. Thirdly, risk
assessment inventories, such as the frequently cited Historical, Clinical and Risk
Management Scales [HCR-20] (Webster, et al., 1997), assess an eclectic mix of
biographical, psychological and environmental risk factors.

The HCR-20 is widely used to assess the risk of violent reoffending, in the USA
(Vitacco et al., 2012), UK (Dolan and Blattner, 2010) and elsewhere. It is
intended as an aid to holistic risk assessment rather than as a numerical
decision-tool. Nevertheless, scores will influence patients’ fates. Follow-up
research (Gray et al., 2004) found that about 75% of discharged forensic mental
health patients whose HCR-20 scores placed them in the higher risk group were
reconvicted for major or minor offences over a three year periodiv, compared with
17% in the lower risk group. Unfortunately, less accuracy was obtained for much
less frequent major offences, the issue of actual concern, than for minor ones.
Gray et al. (2004) concluded that the historical and risk management scales of
the HCR-20 offer moderate predictive accuracy, whilst this and other clinical
scales which assess psychological factors have virtually no prognostic power.
Gray et al. (2007) also found that the clinical (psychological) HCR-20 scale
predicted reoffending more accurately for patients with learning disabilities than
for those with mental disorders. This difference might be associated with whether
patients are capable of presenting themselves strategically in order to manage
perceptions of their riskiness. The issue of informational manoeuvring will be
picked up in the Data Analysis section.

The pattern of findings discussed above demonstrates an inverse relationship
between the probabilistic predictive value of scales and their usefulness for risk
management purposes. Half of the HCR-20 items reference unalterable historical
factorsv. Patients who wish to reduce their measured riskiness, i.e. assessed
probability of reoffending, may need to work extra-hard in order to compensate
for a poor score on this largest sub-scale. The distinction between the HCR-20
clinical and risk management scales is not conceptually clear. But the former
appears to cover mostly the personal factors which forensic and other mental
health services are primarily oriented towards dealing with, whilst the latter mainly addresses the environment which patients will return to. As with historical factors, the patient has hardly any control over the living conditions which they will be sent back to\textsuperscript{vi}. Staff can do little to influence these risk factors on account of organisational fissures between secure and community services (Davies et al., 2006). Moreover, the perceived riskiness of discharged forensic mental health patients can itself impede their community integration (Coffey, 2012).

The predictive power of tools designed to measure the probability of forensic mental health patients reoffending thus appears to be inversely related to their risk management usefulness. The best predictors, derived from recorded history, are therapeutically immutable, and the second best indicators, associated with the return environment, are in practice little easier to modify. Moreover, these risk factors work better prognostically for more frequent minor offences than for the less common major ones which are of primary societal concern. The task at hand for forensic mental health services charged with the rehabilitation of offenders is to identify those who have acquired an acceptably lower probability of future offending despite carrying a troubled history and having to return to criminogenic environments. The inductive prevention paradox makes this task particularly difficult. Consideration of presumed personal risk factors such as ‘negative attitude’ in the HCR-20 brings the analysis back to the question of how practitioners attempt to assess reoffending propensity in a secure environment. The research discussed below aimed to explore patient and provider navigation of this risk assessment task, taking into account reactive processes such as patient attempts to act in ways which lead to them being judged safer and staff efforts to see through presumed self-presentation endeavours.

**METHODOLOGY**

The qualitative data discussed below were drawn from two studies concerned with patient and staff perspectives on risk management for offenders categorised as having mental health problems. A third project completed more recently at a second medium to low security institution in London (Reynolds, 2011) will not be discussed further in this paper. Fieldwork for the first two studies was carried out some time ago, in the period 1999-2003, but similarities with the findings of the third study, completed in 2009, suggests that the micro risk management environment may not have changed much since the first two studies were undertaken. The two studies drawn on in the present paper were located in two UK medium to low secure residential institutions, one catering for offenders with learning disabilities, and the other for those with mental health problems, situated in Northern England and London respectively.

These residential institutions, which offer a range of security levels, can be viewed as ‘risk escalators’ (Heyman, Buswell-Griffiths and Taylor, 2002; Heyman et al., 2004). This organisational form of care is informed by a psycho-logic, not always clearly articulated, in which patients are supposed to travel through
progressively lower levels of security towards discharge as their assessed riskiness is judged to decline in response to therapeutic interventions. Patients can also be sent back up the risk escalator if their progress is considered to have reversed. Such a stepped approach confronts the inductive prevention paradox at each stage, but gives rise to accountability issues most acutely at the point when a patient re-enters the public realm.

Approval from a UK NHS Ethics Committee was obtained for each project. The studies were designed to explore patient and staff perceptions of risk assessment and management in relation to the discharge. Additional information about the two research sites is provided below. Each study was conducted in two stages, with general staff interviews followed by detailed data collection focussed on individual patients. The first stage interviews explored general staff perceptions of risk management within the hospital and in relation to decision-making about discharge. Second stage data collection included, as far as possible, two interviews with selected patients over a 6-20 month period, an interview with a staff member who knew them, and, in the second study, observation of ward rounds and case conferences. Patients were identified through staff and drawn from a range of security levels. Their perspectives do not necessarily represent those of the patient population, but they offered a wide variety of views about residential life and progress towards discharge. Patient interviews lasted 60-90 minutes, and those with staff about 45 minutes.

A grounded theory approach to design, data collection and analysis (Strauss and Corbin, 1998) was adopted, with data collection and analysis undertaken concurrently so that subsequent interviews could take up emergent issues. Transcribed data, mainly interviews, were analysed thematically (open coding), properties of themes and their interrelationships mapped through constant comparison (axial coding), and the analysis organised around a core category (selective coding), the operation of risk escalators, as outlined above. Within this framework, key issues were identified for analysis, including divergent views about the nature of ‘the problem’ (Davies et al., 2006), multidisciplinary teamwork (Shaw et al., 2007) and probabilistic risk assessment, the focus of the present paper. Further details about the two research sites and data collection in each are provided below.

**The forensic learning disabilities study**

Data collection was undertaken during 1999-2000 in an NHS residential facility catering for offenders with learning disabilities. This institution offers a range of security levels from medium secure to unlocked houses. It is located in a rural area of Northern England, several miles away from the nearest town, spreads over a substantial area in its own campus, and has gradually evolved as buildings were added. The campus contains a range of architectural styles including an ultra-modern medium secure unit sealed off by a visually unobtrusive system of electronic locks, family houses formerly in domestic
usage, and low rise pre-1945 brick buildings commonly found in UK asylums. Senior staff members stated that the overall structure with its varied security levels had evolved over time, rather than having been planned.

In the first of the two study phases, interviews were carried out with 13 staff members (two consultant psychiatrists, two clinical psychologists, one forensic service manager, three unit managers, three nurses, one social worker and one occupational worker). No staff member declined to be interviewed. The second study phase involved case studies of 11 patients, nine men and two women, selected randomly from units with different levels of security. As far as possible, case studies included two interviews with each patient and an interview with a staff member who knew them. One additional patient opted not to participate in the initial interview, and two patients could not be re-interviewed. The above patients, and a nurse involved in the care of each (nine nurses in total), were interviewed in a private location on the hospital site. Patients were subsequently re-interviewed in order to explore their perceptions about their progress. The time gap between first and second interviews averaged eight months, with a range of 4-11 months.

The forensic mental health study

The second study was modelled on the one outlined above, and was undertaken in a forensic mental health medium to low security hospital, with data collected in 2000-2003. This institution is located in a deprived inner-city area of London, and has a large proportion of patients and staff from diverse ethnic minorities. The pre-planned facility was originally designed to look municipal rather than prison-like, in keeping with the ideas put forward in the Butler and Reed reports. However, during the mid-2000s, the facility was surrounded by a high mesh fence. A bizarre visual combination was thus created. This shift may reflect changes in the patient population as more serious offenders have been admitted, but may also have been driven by oscillating public attitudes towards mentally disordered offenders.

In the first stage of data collection, 43 interviews with general managers (2), qualified (19) and unqualified (7) nurses, psychologists (3), occupational therapists (3), social workers (3) and doctors (6) were carried out. The sample included senior and frontline staff working in residential and community settings. All but one of the managers and medical staff who were approached consented to be interviewed. Front-line staff were recruited through requests to volunteer, and were therefore self-selected. Nurses on one ward expressed suspicion about the purpose of the project and declined to become involved in the first stage general interviews, although one of these nurses agreed to participate in a case study.

The second study stage involved intensive case studies with 10 patients identified by staff. Data collection included, where possible, two interviews with
each patient, the second undertaken after 11-20 months, an interview with a staff member involved in the individual’s care, and observation of case conferences. Four additional identified patients were not included, two because they declined, and two for clinical reasons. One interviewed patient declined consent for a staff member to be interviewed about his case, and one staff member refused to participate in case studies. Sixteen staff were interviewed about specific patients, providing staff views for nine of the 10 case studies. Two case conferences have been attended and recorded, and five patients have been revisited for a progress update. Of the other five patients, four were discharged during the study period, and one died.

DATA ANALYSIS

The analysis will be presented in two sections: firstly discussing staff and patient perceptions relating to the inductive prevention paradox; and, secondly, considering staff efforts to overcome this problem in relation to patients’ attempts to manage their own assessed riskiness.

The inductive prevention paradox and risk assessment

Data analysis suggests that assessment of the probability of patients reoffending is pervaded by uncertainty. This uncertainty can be contrasted with the official purpose of medium to low secure institutions, articulated in the Butler and Reed reports, discussed above, of providing a setting similar to that of the external world in which risks can be properly assessed. However, one very senior manager, exceptionally, did express confidence in this mission.

We should take somebody who has committed an offence while they have been unwell - bring them in here and be able to - it could be that it is homicide, but bring them in here and treat them, and be able to put them back in the community somewhere around eighteen months to two years.

(General manager, forensic mental health unit)

As illustrated below, staff who worked closer to direct patient care often communicated considerably less confidence that rational decisions about discharge risks could be made. This comparison suggests the hypothesis that those who occupy role positions close to the top of the organisation may, like report and inquiry authors, be more likely to accept the validity of its official mission than those lower down the hierarchy. The latter are engaged with, and accountable for, risk management in individual cases. One ward manager, discussing discharge decision-making believed that ‘it’s more luck than anything’.

I think that we’re such a mixture here of people and patients with difficulties. We’re now dealing with someone with such an extensive forensic history, and such complex needs, that often there’s no clear evidence that things have moved forward. (Ward manager, forensic mental health unit)
The same respondent contrasted the complexity of the processes giving rise to reoffending with the simplification embedded in standardised operational procedures.

*If someone’s worked well within the Home Office [requirements], and had their 12 community trips, and everything has gone according to Home Office plan, but there's still huge anxieties. Because the traits of the personality were, are, still in place, then the doctor will, may well, the team will turn round and say, you know, 'He can go to low secure’ … But he’d done everything by the book … If he has done everything by the book then he will be discharged.* (Ward manager, forensic mental health unit)

This analysis suggests that the forensic health care system responds to the inductive prevention paradox by proceduralising risks which cannot be otherwise assessed. Patients who comply, enabling the appropriate boxes to be ticked, are deemed safe enough for discharge. Such a critique points to a gap between official riskiness assessment and the unknown probability of reoffending. The respondent further argues that uncertainty is mitigated, at least for those making the decision to move a patient down the risk escalator, by the transfer of risk ownership to others, a tactic which attempts to separate risk management decision-making from accountability of subsequent adverse events.

*They would transfer responsibility to going back into the community and to another RMO [responsible medical officer] which, with this particular chap’s history, [would mean] a huge chance he will offend again.* (Ward manager, forensic mental health unit)

As the next quotation suggests, one of the strongest demand characteristics (Orne, 1962) of the forensic mental health care environment is for patient compliance (Davies *et al.*, 2008) which can become confounded with reduced riskiness.

*Patients get worn down really, not really being cared for. But you’re [patients are] beating your head against the wall so many times, so you just accept what’s going on. It’s not really that you become all that better. You’ve just accepted what’s going on … I think it’s just a case of getting used to the environment, or the rules and regulations.* (Charge nurse, forensic mental health unit)

Most patients learnt, sooner or later, to go along with what was required of them. Some, but not all, conflated learning to conform with recovering from their presumed mental disorder, and, by implication, becoming less likely to reoffend after discharge.
Well, I suppose I played the game the right way, you know … That’s to keep quiet and wait, you know, to get better. (Forensic mental health patient)

This patient had become stuck in the system until a nurse ‘frankly’ spelt out what was required.

The nurse told us quite frankly that this [compliance] is the gateway, the doorway to freedom, you know … And I appreciated that, you know. (Forensic mental health patient)

De facto operational reliance on inducing conformity as a means of attempting to reduce riskiness raises two linked issues. Firstly, patients may be discharged simply because they have learnt to meet the expectations of staff who manage secure environments. Compliance can be used as an operational indicator of lowered riskiness, but its association to a reduced probability of reoffending remains conjectural and unexamined. The patient quoted below believed that sex offenders could ‘run’ through the system even though their underlying propensity to offend had not been tackled.

What makes me mad about this place, right, is the fact that, like I say, people running through the system and all that, right … And then you've got, like, on a Saturday, they go down to [local town] by themselves. And owt could happen. Anything's [i.e. children are] around on Saturday. (Forensic learning disabilities patient)

Qualitative research cannot demonstrate the extent to which compliance speeds up release. Nevertheless, this example does illustrate a concern about the validity of risk assessment comparable to that discussed by the ward manager quoted above.

Conversely, patients who do not comply with staff expectations may find that their progress is blocked, even though the relationship between issues arising in this environment and reoffending in the outside world is problematic. For example, consultants expressed alarm at the behaviour of a one forensic mental health patient who had taken hair clippers around other wards without permission, hoping to earn money as an amateur barber. In a community context, this activity might be viewed as commendably entrepreneurial. In the forensic environment it was seen as an indicator of serious riskiness.

A major cause of patients’ actions leading to them being marked out as risky was failure in managing expressions of anger. The significance of such displays depends on their meaningful context, in both everyday life and the forensic environment. The patient quoted below felt that a trivial action had been wrongly classified as a riskiness indicator.
I walked across to the table, the pool table, tapped on the top of it, and she [the nurse] wrote down that I was feeling aggressive, and that, and all things like that. And I just thought, ‘Well, one tap on the table’. I thought that was entirely wrong. So I said. She discussed it. It came out in the ward round that, she wrote that, which was wrong, out of order. (Forensic mental health patient)

This patient felt that he had merely been indicating that his right to a turn at the pool table was being violated. The expansion ‘and all things like that’ conveys a sense that an edifice of risk reasoning was being built on a misreading of a single observation about an action which would not register as a risk indicator elsewhere. By overtly challenging this interpretation, the patient may have further harmed his discharge prospects. A propensity for mental health professionals to interpret everyday behaviours in terms of pathological labels (Rosenhan, 1973) is well-documented, but takes on new forms within a risk assessment framework. Patients and staff occupy an enclosed, highly frustrating environment in which interpersonal conflict may be expected. In addition, patients may be subjected to minute observation designed to determine the probability of them reoffending. This combination of close confinement and total risk assessment may obscure the issue of primary concern, namely the likelihood of a patient harming others after discharge.

A final illustrative example of contested risk assessment is of particular interest because it illustrates how a patient’s direct statement of an intention to offend can be discounted within a medical interpretive framework.

He [patient] said that he wanted to [commit serious offences]. And they still let him go because he turned round and said, ‘Well, I made it all up. I just wanted to go and see my mum’…. So, you know, as a nursing team, the day before, we had sat around just kind of gob-smacked that the consultant had said that he could go … I would not escort him. (Health care assistant, forensic mental health unit)

This patient’s tactic may have worked because the consultant viewed his behaviour as symptomatic of illness, and ‘prescribed’ a home visit. When asked why the consultant had agreed to his parole, a decision which the above respondent depicted as ‘complete madness’, she cited the reason given in the patient’s medical notes, namely ‘to allay his [patient’s] anxiety’. Outside forensic mental health settings, people often make threats, such as ‘I will kill you’ which are not intended or taken literally. The forensic context frames such statement as potentially threatening. Health professionals are faced with the task of differentiating serious statements of intent from merely metaphorical threats. In this case, doctors felt that they knew the patient well enough to rule out real risk of offending. However, if the patient had harmed someone whilst on leave, they would have been held to account by judges of responsible risk-taking.
Strategies for managing the inductive prevention paradox

Some staff and patients expressed concern about the problematic validity of risk assessments, as illustrated above. In general, the data suggest that, not surprisingly, staff gave more attention to reducing the risk of released patients reoffending than they did to that of detaining patients unnecessarily. Three strategies which staff adopted in order to detect presumed underlying personal riskiness are discussed below: discounting good behaviour; mini-trials; and formalised risk assessment.

Discounting good behaviour

Given that patients mostly want to be released as early as possible, they might be expected to attempt to act in ways which would reduce their assessed riskiness. Staff, in turn, might try to see through such attempts at self-presentation in order to minimise the risk of released patients reoffending. One forensic learning disabilities worker, discussing this issue, said that male patients would be asked questions designed to test their truthfulness such as whether they would look at a woman with large breasts! Staff described covertly observing a forensic learning disabilities patient who had committed offences involving children when he went to a swimming pool in order to see if he showed an inappropriate interest in them. This approach, an informal version of personality test 'lie scales', provides an obviously fragile method for checking patient veracity. A psychiatrist indicated that staff might deliberately withhold revealing the purpose behind an activity involving a patient so that it could be used as a test of their inclination to reoffend. The ‘star patient’, discussed next, was seen as operating a policy designed to make his mental health appear as good as possible, and thereby maximise his prospects of early release.

Every time I stop, ‘Oh I’m fine, I’m alright’. ‘Have you got anything you are worried about?’ ‘No.’ … He’s all pleasant. He looks normal … We know he is the ‘star patient’ and everything, but [laughs] we have to watch him, [given] what he did before, you know. (Nurse, forensic mental health unit)

The patient’s conduct was viewed as too good to be true. This suspicion was framed by awareness of the seriousness of his previous offending, illustrating the difficulty which patients experienced about compensating for historical risk indicators, as discussed above in relation to risk assessment tools. Patients faced with this bind may adopt a more subtle approach, first ‘faking bad’ so that their subsequent conversion to low riskiness might appear more credible. Mulvey, the fictitious nineteenth century prisoner described in the opening quotation from Star of the Sea employed just this tactic.

Mini-trials
Testing through mini-trials involves allowing a small temporary increase in autonomy in the hope of assessing the probability of future reoffending more accurately, whilst limiting the risk of it actually occurring. Success can be built on progressively, for instance by lengthening parole periods. This strategy had been adopted with the ‘star patient’ discussed above.

**Nurse:** *I personally think, when he ['star patient'] goes out on parole, that's a big test for him, because he goes out on a Saturday to [large town], and [large town] is quite far, and anything can happen then … If something really pushed him, he would do something.* (Nurse, forensic mental health unit)

The quotation brings out the two features of mini-trials mentioned above, that they provide a somewhat more realistic test of riskiness than can be managed in the secure environment, but only at the price of accepting a chance that the patient might reoffend during the test period. The strategy can be compared to that traditionally adopted for inductively testing the safety of novel foods by eating progressively larger amounts. Its limitations can easily be identified, and will be considered further in the Discussion. As the above respondent indicates, a risk of immediate disaster is inescapably incurred. Conversely, patients might conceal their offending proclivity until permanently discharged.

**Formalised risk assessment**

Tools such as the HCR-20 which are used to assess reoffending risk can also take on a symbolic function. This way of ‘managing’ risk assessment is illustrated by the following extract derived from observation of a ward round. The consultant quoted made the comment cited below after a lengthy discussion concerning difficulties arising from the actions of a female patient, including conflicts about bathing and money matters, and accusations directed at male staff.

**Consultant:** *What risk assessment was done. The pink thing [risk assessment form]? Do we need to assign numbers?* [Senior house officer reads out numbers from the ward round summary.] *Make a point of noting risk to others on the ward round minutes and notes. We need to be vigilant.* [Moves on to next patient.] (Ward round, forensic mental health unit)

The reading out of probability ‘numbers’ combined with an admonition to be ‘vigilant’ conveys a sense that an authoritative, scientifically rational resolution has been achieved, even though it skirts over their problematic meaning.

Measuring patient progress towards a level of riskiness which would justify their discharge relies on observing their behaviour. However, many of the issues which arise in a secure setting bear little relation to those which would be of concern outside and *visa versa*. The validity of the whole process depends upon assuming either that patients are ‘judgemental dopes’ (Garfinkel, 1967) who are presumed not to take into account the anticipated implications of their actions for
risk assessment, or that their self-presentation strategies can be penetrated through information games or psychometrics. The information management issues identified by Garfinkel and Goffman (1959/1970) half a century ago have now taken on new forms in relation to probabilistic risk assessment.

**DISCUSSION**

It has been argued in this paper that the rationality of risk assessment is undermined by the operation of the inductive prevention paradox. In relation to the specific example discussed, forensic mental health service providers are expected to assess the probability of a patient harming others in the future after discharge by obtaining observational ‘evidence’ in a present environment designed to prevent just such events from occurring. They are required to answer the question of how, if at all, a patient’s riskiness has been changed by long-term incarceration, even though the inductive prevention paradox cuts off the supply of ‘evidence and rigorous analysis’ (South West London Strategic Health Authority, 2006, p.9) which those who conduct retrospective inquiries often assume to be readily available. As Warner (2006) has argued, inquiries have acquired the function of modern allegories, precautionary tales which fuel organisational defensive practices oriented towards minimising reputational risks (Rothstein, 2006). Their role as risk owners puts staff into a bind which they struggle to manage. They lack an inductive evidence base for probability estimation, but risk condemnation if a discharged patient seriously reoffends.

The ultimate source of this bind is the prevailing cultural assumption that science can banish risk. This unfulfillable promise (Hansson, 1993) comes up against a double limitation: that probabilistic reasoning does not preclude the occurrence of adverse events; and that prevention blocks out risk managers’ views of the most relevant evidence. Official probabilistic reasoning based on induction from observed rates entails heuristic acceptance of the ecological fallacy (Rose, 1981). This approach to prediction requires the assumption to be made that each member of a constructed category or sub-category ‘carries’ its aggregate proportion of adverse events, (and the further presupposition that such historic rates provide a good guide to the future). But risk managers considering offender discharge are expected to make correct predictions in individual cases rather than for patient categories, somehow allowing for preventative measures currently in place and discounting patients’ attempts to influence their risk assessment (Dixon, 2012).

The inductive prevention paradox can be escaped from if proxy measures associated with the probability of the outcome of concern, but unaffected by the operating preventative measures, can be identified. Unfortunately, the prognostic need for such measures does is not matched by their availability. The limitations of the ‘tools’ available for risk assessment in a forensic mental health context were reviewed in the Introduction. It was noted that the best predictor, offending history, adds no predictive value with respect to changes in the probability of a
patient reoffending; that staff in practice have little sight or control of the post-discharge environment, the second-best predictor; and that the psychological factors which forensic mental health services are oriented to changing offer no predictive value⁸. In practice, peaceably serving time as a patient may serve as the main de facto approach to risk reduction, as articulated by the patient, quoted above, who had taken on board the advice ‘to keep quiet and wait, you know, to get better’⁹.

The efforts of forensic mental health service providers to compensate for this prognostic limitation were obviously flawed. Their strategies included trying to see through patient self-presentations, using mini-trials, and relying on the illusory precision of numerical risk assessments. Although the information games required for the first two tactics were depicted long ago, the inmates of the asylums observed by Goffman (1961/1991) were not subjected to constant risk assessment as they would be today. In the heyday of asylums, staff stripped ‘dangerous’ patients (Castel, 1991) of identities which they considered defective in order to rebuild them. This form of governmentality has been replaced by one framed in terms of risk. Patients are sifted through a probabilistic filter which is designed to allow only acceptably safe individuals to pass through to the world outside. Adoption of this interpretive framework can lead to the creation of new forms of bureaucratically driven risk-blindness (Godin et al., 2006) underpinned by faith in probability ‘numbers’ generated by ‘tools’. Mini-trials have the merit of being tailored to the specific risks which concern staff, and may offer some prognostic utility with respect to countering the obscuring effect of the inductive prevention paradox. However, their limitations must not be lost sight of. Patients who are determined to reoffend may do so during the brief window of opportunity offered by a mini-trial, consciously bide their time until monitoring is relaxed in response to a successful mini-trial, or find that they cannot refrain from reoffending over a longer time-frame.

The inductive prevention paradox is particularly likely to trap decision-making in an evidence-impoverished virtual bubble when the contingency of concern is culturally abhorred, as in the case of ‘mad’ or ‘dirty’ offences against the person, concern about which is fuelled to a greater or lesser extent by media amplification in different countries. However, other more or less emotionally charged examples can easily be identified. The most direct comparison is with prisoners attempting to negotiate release via parole boards who report similar concerns about informational game-playing to those illustrated in the present paper (Muhammad, 1996).

More broadly, anyone who has taken preventative measures faces the question of what damage might result from lifting them. Although prophylaxis can be resumed if necessary, its reapplication might not erase irreversible consequences of temporary withdrawal. For example, a frail older person who has decided that going out alone is too risky might suffer a serious injury if they tried to test their current capabilities. Similarly, a person living with depression
might commit suicide if they stopped taking medication; and a Crone’s disease sufferer might experience uncontrollable flare-ups if they abandoned anti-inflammatory drugs for a trial period. As with incarcerated mental health patients, preventative measures may themselves change risk conditions in ways which are hard to assess. An older person who has confined themselves to their home for a lengthy period might become more vulnerable to falls through lack of practice if they do eventually go out. Similarly, the long-term anti-depressant user might experience withdrawal symptoms from these addictive substances, or suffer long-term psycho-physiological damage. Unless they remove protective measures, thereby facing the risk of irreversible damage, individuals in these predicaments cannot know whether they are still necessary or not.

CONCLUSION

Policy-makers and practitioners are entitled to ask how they should be expected to respond to the difficulties arising from the inductive prevention paradox. Four constructive suggestions can be drawn from the present analysis, and are outlined briefly below along with the formidable barriers to their achievement.

Firstly, societal expectations about the potential of risk assessment in this and other domains in which probabilistic inductive inference is obscured by prophylaxis could be lowered. In theory, the prevailing de facto blame culture could be replaced by one which allows for positive risk-taking (Titterton, 2005). However, as noted, mass media approaches to crime which fuel defensive practices appear to be associated with wider structural factors such as fragmented media ownership and conflictual politics which cannot easily be modified (Walgrave and Sadicaris, 2009). Secondly, patients could be actively engaged in an open risk assessment process (Langan and Lindow, 2004), rather than being deceived in order to try to prevent them from concealing their ‘real’ riskiness. This approach would require a substantial shift towards trusting patients as ‘service-users’, and might not allow those who are determined to conceal their intention to re-offend to be identified. But it would, at best, enable patients and staff to collaborate rather than play risk assessment information games which rely on concealment. Thirdly, top-down moralising about poor inter-agency and multi-professional collaboration could be replaced by serious attempts to analyse the organisational reasons why such failures endlessly recur, particularly the often underestimated barriers to multidisciplinary collaboration (Shaw et al., 2007). Risk systems theory (Japp and Kusche, 2008) provides one useful starting point for analysing such barriers by postulating that organised groups unreflectively construct and orient to distinctive risk objects. Finally, serious attempts to address the environments which offenders return to (Mullen, 2000; Brett et al., 2007; Coffey, 2012) would complement the present focus on assessing and reducing personal riskiness. Although currently inhibited by economic, organisational and political barriers, multidisciplinary attention to the quality of the lives which forensic mental health patients will return to perhaps
offers the most promising but also challenging approach to their safer rehabilitation.

REFERENCES


The 'inductive prevention paradox' should be distinguished from the widely discussed 'prevention paradox' (Rose, 1981) which arises from attributing a risk associated with an aggregate category such as alcohol consumption to individuals who meet the specified criteria for category membership. Both paradoxes are bound up with the limitations of probabilistic thinking. But the latter, which could be termed the 'ecological prevention paradox', derives from the requirement to shift between the aggregate and the individual in order to quantify probabilities (Heyman et al., 1998; Hunt, 2003). The inductive prevention paradox results from limitations in the observational evidence base in individual cases.

The extent and nature of media crime coverage varies across countries. One comparative study of 11 countries (Walgrave and Sadicaris, 2009) concluded that the proportion of national TV news-time devoted to this topic was positively associated with higher crime rates, greater media fragmentation and competition, and stronger political polarisation. The UK was found to have some of the highest percentages of TV news items concerned with crime, nearly double those for France, together with a particularly strong slant towards personalistic accounts of violent offending. (The USA had an unexpectedly low rate of crime coverage which, the authors suggest, may have resulted from selecting national rather than more frequently used local TV stations.) A febrile media culture, reinforced by the predominance of crime as a TV drama theme and political competition to be 'tough on crime', leaves UK forensic mental health services facing a particularly difficult risk management dilemma.

Tools direct selective service attention to particular adverse events, in this case the risk of reoffending as against, for example, becoming depressed or being attacked by members of the community. They thereby carry implicit value judgements (Heyman, 2012).

Probabilities can only be quantified in relation to a temporal horizon, in this case three years, beyond which adverse events are not taken into account. Practically focused risk managers tend to frame time unreflectively.

More accurately, the probability of a patient reoffending is related to recorded history. Patients can influence their 'history' in this sense, for example by concealing previous offending. However, once their offending history has been encoded in a patient record, patients are cannot change it, unless they can demonstrate their innocence, a very unlikely possibility for those who have committed offences against the person.

Discharged patients may be directed away from localities associated with former offending. However, they will thereby also be separated from social networks and familiar surroundings. They tend to be discharged into areas of serious socio-economic deprivation which are associated with additional problems such as high crime rates and drug problems.

The concept of risk ownership originated in corporate governance where it is used to convey a top-down model of social order in which a senior manager at board level is made accountable for each risk which the organisation is deemed to face. As with the idea of risk management, the notion can be applied more generally to everyday life. For example, in the forensic mental health sphere, close relatives may decline to take back responsibility for a discharged offender (Heyman et al., 2010, p. 34).

A parallel debate has taken place in the criminal justice system, with advocacy of a 'tool' for assessing reoffending risk, the Psychopathy Checklist–Revised (PCL-R) challenged by sceptics (Gendreau, Goggin and Smith, 2002).

As this example illustrates, the 'lens of risk' may give a new frame for interpreting the social organisation of patient hospital stays (Roth, 1963), in which the passage of time is seen to somehow reduce riskiness.
Randomised controlled trials (RCTs) can provide counterfactual evidence at an aggregate level. For instance, patients in the placebo group can be expected to have done as well on average as those in the treatment group if they had received an experimental drug. However, this aggregated knowledge offers limited predictive accuracy in individual cases, and the methodology of RCTs is difficult or impossible to apply with respect to long-term, complex interventions.