Objective Predictors of Outcome in Forensic Mental Health Services – A Systematic Review

Ottilie Sedgwick 1,2,3, Susan Young 2,4, Mrigendra Das 2 and Veena Kumari 1,3

1. Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King’s College London, London, UK

2. Broadmoor Hospital, West London Mental Health Trust, Berkshire, UK.

3. National Institute for Health Research Biomedical Research Centre for Mental Health, Institute of Psychiatry and South London and Maudsley NHS Trust, London, UK

4. Centre for Mental Health, Faculty of Medicine, Imperial College London, London, UK.

Correspondence to:
Ottilie Sedgwick
Department of Psychology (PO78)
Institute of Psychiatry, Psychology & Neuroscience
De Crespigny Park
London SE5 8AF
Email: ottilie.sedgwick@kcl.ac.uk
Ottilie Sedgwick (MSc) is a PhD student at the Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London.

Professor Susan Young (PhD, DClinPsy) is a senior lecturer at Imperial College London, an honorary consultant clinical psychologist at Broadmoor High Secure Hospital, and the director of forensic research and development at West London Mental Health Trust.

Dr Mrigendra Das (MBBS, MRCPsych) is a consultant forensic psychiatrist at Broadmoor High Secure Hospital, West London Mental Health Trust.

Professor Veena Kumari (PhD) is a professor of experimental psychology at the Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King’s College London.
Abstract

This systematic review aimed to examine whether neurobiological methods, or other methods independent of clinical judgement, have been investigated to assist decision-making in forensic mental health services, and if so, whether this may be a useful strategy for predicting outcomes. OVID-Medline, Embase and PsychInfo (inception-January 2015) were searched, limiting to English and human studies, using terms relating to “predict”, “outcome”, “psychiatry” and “forensic” to identify primary research articles reporting on predictors of outcome in forensic mental health services not reliant on clinical judgement/self-report. 50 studies investigating demographic, neuropsychological/neurophysiological and biological predictors were identified, reporting on three, broad outcomes: i) inpatient violence, ii) length of stay, iii) reoffending. Factors associated positively, negatively, and showing no relationship with each outcome were extracted and compiled across studies. Of various demographic predictors examined, the most consistent associations were between previous psychiatric admissions and inpatient violence; a more ‘severe’ offence and a longer length of stay; and young age and reoffending. Poor performance on tests of cognitive-control and social-cognition predicted inpatient violence while a neurophysiological measure of impulsivity showed utility predicting reoffending. Serum cholesterol and creatine kinase emerged as biological factors with potential to predict future inpatient violence. Research in this field is in its infancy, but investigations conducted to date indicate that using objective markers is a promising strategy to predict clinically significant outcomes.
Introduction

Outcomes in forensic mental health services are varied and often poor. In 2007 around 50% of patients detained under the legal category ‘psychopathic disorder’ in the United Kingdom had a stay in hospital exceeding ten years.1 Lengthy admissions were also identified in one German study finding that some patients stayed as long as 43 years.2 Further, prospective follow up studies of discharged mentally disordered offenders (MDOs) have shown a relatively high rate of reoffending, with one in eight men being convicted for another grave offence after discharge from medium security services in the UK.3 This has significant implications in terms of public protection, cost to the taxpayer, and the ethical position of detaining individuals for treatment which may not be efficacious.

Current methods of predicting outcome include a multidisciplinary assessment of need (i.e. criminogenic and clinical factors which require intervention) often involving the use of structured professional judgement instruments to assess the level of risk, generally in the context of treatment planning4,5. The Historical Clinical Risk Management (HCR-20)6 scheme is an example of this, and has shown good predictive validity for future violence7. The psychopathy checklist (PCL-R) has grown in popularity as a quasi-risk-assessment tool due to the demonstrated link between high PCL-R scores and both inpatient violence and community reoffending8,9. However, while these assessment tools supersede unstructured clinical decision making10, they still rely on clinical judgement/decision making to draw conclusions. This is particularly relevant when considering the forensic population, many of whom are diagnosed with disorders which are characterised by deceptive behaviours (e.g. antisocial personality disorder, taken from DSM-5; “Deceitfulness: dishonesty and fraudulence; misrepresentation of self; embellishment or fabrication when relating events.”). Further, it is plausible that offenders may wish to present as low risk in order to secure early discharge, adding a further complication for clinicians making assessments of need.

A recent review11 has identified wide variation in the rate of violence observed from those who are classified as ‘high-risk’ using nine of the most widely used risk assessment tools, both within and between risk assessment schemes. When considering this alongside evidence suggesting that there is very little change in HCR-20 scores across an individual’s stay in high-security hospital despite them engaging in risk-focussed treatment, it calls into question the clinical utility and sensitivity of such tools. For example, Morrissey and colleagues12 found a change of one point or less (possible score range 0-10) in the dynamic clinical and risk scales, across five years. Although the clinical scale scores were significantly lower in the group about to be discharged compared to those who were still
resident in the hospital, the risk scale and total scores were comparable, suggesting that these scales are either not sensitive enough to capture a reduction in risk, or that clinicians are not regarding this information as useful in their decision making about discharge.

A growing body of evidence has shown that an array of neurobiological factors are associated with violent behaviours in mentally disordered populations\textsuperscript{13-17}, and it may be that some of these correlates could also assist clinicians working in forensic services to make decisions about treatment planning, risk and discharge. Such factors, which are objective and measurable, reduce the likelihood of errors of judgement being made. Consideration of these factors alongside methods already employed could enhance the amount of information available, and thus potentially improve decision making or identify areas of outstanding need. This could theoretically lead to improved outcomes for patients, the public and the taxpayer, via more appropriate treatments being offered, fewer premature discharges and more efficient services, respectively.

This systematic review aimed to identify and evaluate studies which have assessed objective predictors of outcome in forensic mental health services (i.e. did not rely on self-report or clinical judgement) to gain a perspective on how far these correlates have been used by the scientific and clinical community, and to assess the potential usefulness of such markers in further research and subsequently in clinical practice.

**Method**

OVID-Medline, Embase and PsychInfo (inception-January 2015) databases were searched using the following four terms combined with AND:

1. predict* OR prognos* or marker
2. outcome OR length of stay OR duration of stay OR length of hospitalization OR duration of hospitalization OR reoffenc* OR recidiv* OR violen* OR function*
3. mental disorder OR psychiatr* OR mental ill*
4. forensic OR secur* OR incarcerated
A screen of the results for relevance was then conducted on a title/abstract basis. If insufficient information was given in the abstract, the full text was retrieved before making a decision. Studies were assessed for inclusion against the following criteria:

1. All participants were MDOs admitted to inpatient forensic psychiatric services. For the purposes of this review, an MDO is defined as an offender with a diagnosed mental disorder, who is deemed to require treatment in psychiatric services. Individuals residing in prison who have a mental disorder were not included as it is highly likely that individuals who are deemed treatable within prison (as opposed to secure psychiatric hospitals) are qualitatively different. Further, ‘specialist’ offender groups (adolescents, e.g. 18, learning disability, e.g. 19) were excluded to keep the study samples as homogeneous as possible.

2. Studies which included an objective predictor of outcome (as defined as a factor which does not rely on clinical judgement or self-report, e.g. biological, neuropsychological, demographic factors), with outcome defined as one of the following: length of stay, violent incidents (inpatient or community), functioning, clinician rated risk/need.

3. Only primary research articles with an abstract were included (e.g. not theses, reviews). The reference lists of relevant reviews were examined to identify any papers not returned by the initial search.

4. Studies were only included if they used a prospective, or pseudo-prospective, design (i.e. looking forward over time) to assess predictive ability. Studies which reported on the ability of static (i.e. demographic) factors to predict outcome were also included; these did not necessarily need to be prospective as static factors by definition are temporally stable.

5. Studies were excluded if they were reviewing the predictive validity of risk assessment tools. This literature is large and robust and has been reviewed elsewhere e.g. 20-22. Further, these tools require the assessment of a combination of demographic and clinical factors which may relate to risk collectively, but often individual item predictive validity is not given.

6. Articles referring solely to competency to stand trial were also excluded. This intervention involves treating the underlying disorder and educating the individual about the American legal system so
that they are able to stand trial—it is not analogous with the typical treatment MDOs receive (i.e. the focus is to restore competency).

Data Extraction

For each study, predictors associated positively with the outcome variable of interest (e.g. associated with an increased likelihood of violence), predictors with a negative association (e.g. associated with a decreased likelihood of violence) and examined variables with no relationship (e.g. no relationship to violence) were extracted. Studies were examined and any factors identified by the authors as ‘statistically significant’ were extracted. This included significant differences between relevant groups (e.g. between reoffenders and non-reoffenders) and significant positive or negative predictors of outcome. Variables that were examined by the authors but had no significant effects were included in the ‘no relationship’ category.

Predictor variables were then compiled into a spreadsheet, and studies which reported on the same broad predictors for the outcome of interest were recorded. Categories which were conceptually similar but perhaps not described in the exact same terms (for example ‘severity of offence’ and ‘a violent or homicide offence’) were combined to reduce the number of discrete predictors.

Results

The search returned 1896 results. See Figure 1 for the flowchart of study selection.

*****Figure 1 about here*****

50 articles were retained in the final review which included data on objective predictors of outcome in forensic mental health services. Studies were categorised into three, broad outcome groups, those reporting on predictors of: 1) inpatient violence, 2) length of stay in forensic inpatient services, and 3) community reoffending. Further, the types of predictor could also be delineated into three categories. These were i) demographic (42 studies), ii) neuropsychological/ neurophysiological (4 studies) and iii) biological (4 studies) predictors. The term “demographic” is used here as a broad, all-encompassing term to refer to static, historical factors, including clinical, offence-related, developmental, institutional and sociodemographic factors.
Predictors of Inpatient Violence

a) Demographic Predictors

38 separate demographic factors across eight studies\textsuperscript{24-31} were identified as predictors of inpatient violence (Table 1). Of these, 16 factors were considered in more than one study. Only one factor, previous psychiatric admissions, was found to be associated with inpatient violence in the majority of studies which examined it; two studies found a positive relationship between number of previous psychiatric admissions and inpatient violence\textsuperscript{24, 31}, whereas one study found a null effect\textsuperscript{28}. One of these studies assessed seclusion episodes as opposed to inpatient violence directly\textsuperscript{31}; however, all seclusion incidents were related to aggressive behaviour, apart from one episode of self-harm.

Another demographic factor, young age, was examined by six studies, of which three found a positive association\textsuperscript{27, 29, 31} and three found no association\textsuperscript{24, 25, 28}. Similarly, a history of violence was found to be associated \textit{with inpatient violence} in two studies\textsuperscript{24, 28}, and not associated in three studies\textsuperscript{29-31}.

Other factors examined by two or more studies and found to be unrelated to inpatient violence are listed in Figure 2. Notably, a history of substance use\textsuperscript{27, 28, 30, 31}, diagnosis\textsuperscript{24, 31} and gender\textsuperscript{24, 28, 30, 31} did not emerge as consistent predictors across studies (Figure 2).

\textit{*****Table 1 and Figure 2 about here*****}

b) Neuropsychological Predictors

One study\textsuperscript{32} reported the ability of neuropsychological assessments to predict aggression amongst 23 male forensic inpatients (n=16 with a principal diagnosis of schizophrenia). Aggressive behaviour was monitored over the year following testing using the Overt Aggression Scale\textsuperscript{33}. The results demonstrated that poor visuospatial processing \textit{[assessed by the Judgement of Line Orientation Test (JLOT)\textsuperscript{34}]}, poor cognitive inhibition \textit{[scores on the Stroop Colour/Word Test (SCWT)\textsuperscript{35}]} and the number of misperceptions of an angry voice in an emotional recognition test could reliably predict the frequency of subsequent aggression. Scores from the JLOT and SCWT were also significantly correlated with the severity of aggression.

A similar study\textsuperscript{36} reported a five week follow-up of ten forensic inpatients. Contrary to expectation, performance on a measure of behavioural inhibition (the Stop Task\textsuperscript{37}) was better at a trend level amongst those who were involved in aggressive incidents compared to those who were not, suggesting that those who were more impulsive were
involved in fewer incidents. However, this study was significantly limited by its small sample size and low rate of recorded incidents (12 incidents, conducted by five patients), and thus the results must be interpreted with caution. In addition, no information regarding diagnosis is given by this study, leaving questions as to the generalizability of the results to other populations.

A further study examined clinical outcome, need and risk in high-security hospital, which are all facets sensitive to inpatient violence. Thirty newly admitted men with schizophrenia were assessed on a number of neuropsychological tasks including an assessment of IQ, processing speed and working memory using the Wechsler Adult Intelligence Scales, in addition to the Trail Making Test and the SCWT. Further, two social cognitive tasks were conducted, the Revised Eyes Task and a Modified Advanced Theory of Mind Test. Outcome measures included the Health of the Nation Scales – Secure version (HoNOS), the Camberwell Assessment of Need – forensic version (CANFOR) and the HCR-20, assessing clinical, social and functional outcome, need and risk respectively, at three year follow-up. Although a number of non-social cognitive tasks showed utility in predicting some outcomes of interest (e.g. Trail Making part B was significantly correlated with scales from the HoNOS, the total CANFOR score and HCR-20 risk management scale), the overwhelmingly most predictive test was the Revised Eyes Test. After controlling for all other variables, the Revised Eyes Test score could significantly predict total CANFOR score, the risk management score on the HCR-20 and the social scale score of the HoNOS.

Thus, patients with schizophrenia who were less able to interpret emotional information from the eyes were likely to have higher ratings of unmet need, poorer social functioning and a higher level of assessed risk. This may be relevant to the Violence Inhibition Mechanism theory, according to which poor interpretation of negative facial expression removes inhibitory influences which serve to stop violent behaviour through negative reinforcement of the unwanted (aggressive) behaviour. Poor theory of mind may also reduce the capacity for cognitive empathy, or understanding typical social rules which could lead to social conflict and potentially violent behaviour.

Finally, one demographic study extracted evidence of “cognitive impairment” (present/absent) from patient files, and found that this was a significant predictor of frequent violent behaviour amongst inpatients. Although there is not detailed explanation of the nature or severity of cognitive impairment in these participants, this study supports the assertion that cognitive dysfunction may be related to aggressive behaviours.

\textit{c) Biological Predictors}
Four studies\textsuperscript{46-49} examined biological predictors of inpatient violence, although three of these studies\textsuperscript{46, 47, 49} were conducted within the same sample. Two \textsuperscript{46, 48} related to serum cholesterol levels, while two\textsuperscript{47, 49} were concerned with creatine kinase elevations.

\textit{i) Serum Cholesterol}

The serum-cholesterol level of 106 forensic inpatients at admission was examined, and subsequent aggressive incidents towards others or themselves over the following two years (pseudo-prospective review of medical records) was followed up\textsuperscript{46}. The sample was divided into high (\(\geq 200\) mg/dl) and low (<200 mg/dl) cholesterol groups, and the difference in aggressive incidents (frequency, severity and type) was investigated. While the two groups did not differ with regards to severity or type of aggression, the frequency of aggression in the low cholesterol group was significantly increased. Interestingly, the relationship between cholesterol level and frequency of aggression was non-linear, with aggression being most frequent within the range 160-170 mg/dl.

A similar investigation\textsuperscript{48} was conducted aiming to determine an optimum cut-off point for predicting aggression using serum cholesterol levels. Using male participants detained in forensic hospital, the sample was divided into those who had been secluded at least once over a 28 month period (n=195) and those who had not been secluded (n=202). When comparing these groups, the secluded group had significantly lower total serum cholesterol. Using receiver operating characteristic analysis, the optimum cut-off for predicting those who would be secluded for any reason was 5.3 mmol/l. However, for patients who spent a longer duration of their detention in seclusion for aggression/self-harm, perhaps considered the most frequently aggressive patients, the optimum cut off was 4.3 mmol/l. When converted into mg/dl (as used in the previous investigation) this equates to approximately 165 mg/dl, which is highly consistent with the 160-170 mg/dl range found for the most frequently violent patients in the aforementioned study\textsuperscript{46}. The difference in cholesterol level between the two groups was independent of body mass index and medication.

\textit{ii) Creatine Kinase}

One study\textsuperscript{47} investigated the predictive utility of creatine kinase (CK) as a marker of aggressive behaviour in 164 male forensic inpatients, again using a pseudo-prospective design. CK is an enzyme involved in in-situ energy production in cells\textsuperscript{50}. The sample was divided into high or low aggression, based on a median split procedure on scores for the severity, frequency and type of violence as determined by the Overt Aggression Scale (verbal vs. physical). In all three comparisons (severity, frequency and type), the CK levels were significantly higher in those
who were more frequently violent, engaged in more severe violence, and in those who used physical as opposed to verbal aggression.

An association between assaultiveness and use of restraint prior to CK levels being determined was also observed. Those who had been assaultive during their admission and those who had been restrained had higher observed CK levels than those who had not. Importantly, a significant interaction between these factors was observed, in that those who were assaultive/restrained and then engaged in subsequent violence had significantly increased CK levels (around a five-fold increase) during their admission compared to those who were assaultive/restrained and then not violent. This suggests that, of those patients who present management problems during their admission, the likelihood of subsequent aggression can be gauged by assessing CK levels. These findings were irrespective of diagnosis, recent physical exercise, recent accidents or recent intra-muscular medication. However, two caveats were noted: 1) these findings were only significant in those patients taking antipsychotic medication, and 2) CK levels were not sensitive to change in aggression – they did not increase prior to an aggressive incident, nor decrease afterwards. Despite this, the authors assert that using a >200 U/l cut-off could correctly predict future assaults in 94% of cases, compared to using prior assaultiveness alone as a predictor (64%).

A further study on the same sample examined CK as a function of ethnicity and aggression. While the results demonstrated that CK levels were higher in African Americans than in Caucasians, and that African Americans were more likely to be physically aggressive compared to Caucasians, the increased levels of CK observed in African Americans was still significant even when the effect of aggression was covaried out.

Predictors of Length of Stay

a) Demographic Predictors

A total of 44 diverse predictors were examined in relation to length of stay, with 25 of these being examined by more than one study (Figure 3). The factor which most studies examined was severity of offence. Unsurprisingly, nine, out of ten studies found that a more ‘severe’ offence was related to a longer length of stay. This is supported by two studies examining the effect of a restriction order on length of stay (administered to patients in the UK who are considered to be particularly high-risk), which both showed a lengthening effect. Three studies found that having a psychotic disorder was associated with a longer length of stay, although one study found the opposite (shorter stay), and one found no significant effect. In addition, three studies found no effect
for ‘diagnosis’ on length of stay (which included psychosis)\textsuperscript{52, 55, 56}, however, it is notable that in two of these studies there was a very small proportion of offenders not diagnosed with a psychotic illness, suggesting limited sensitivity to find an effect. Two out of three studies which examined absconding during hospitalisation found that this was associated with a longer stay\textsuperscript{2, 61}.

Previous offences was found to be unrelated to length of stay in all six studies which examined this\textsuperscript{2, 51, 52, 55-57}, providing strong evidence that it is the severity, as opposed to the extent, of offending which is implicated in how long MDOs remain in services. Other examined factors for which no clear association emerged are detailed in Figure 3 and Supplementary Table 2.

\textit{b) Neuropsychological/Neurophysiological and Biological Predictors}

No studies examining the effect of neuropsychological/neurophysiological or biological variables on length of stay were identified.

******* Figure 3 about here*******

Predictors of Community Reoffending

\textit{a) Demographic Predictors}

Community reoffending, encapsulating re-arrest, readmission, recidivism etc., was the outcome of interest in the majority of the papers (n=25). Again, a large and diverse number of factors (total 66) were considered across studies (Table 2), with 27 factors only considered in a single study. The most frequently examined predictor was previous offending, examined by 18 studies\textsuperscript{3, 52, 54, 55, 61-74}. 67\% of studies examining previous offending found an association with reoffending. Young age at admission or discharge was investigated in 15 studies\textsuperscript{3, 26, 52, 54, 55, 62-67, 69, 73, 75, 76} with 67\% finding a positive effect, while the effect of a shorter length of stay was examined in 12 studies\textsuperscript{3, 55, 62, 66-68, 70, 73-77} and 50\% found a it was associated with reoffending.

Male gender\textsuperscript{3, 52, 55, 63, 64, 67, 69, 73, 76, 77}, race\textsuperscript{3, 52, 55, 62, 65-67, 69, 76, 78} and being single\textsuperscript{3, 54, 62, 65, 69-72, 75, 77} were investigated in 10 studies each, with positive findings indicated in 40\%, 20\% and 30\% of studies, respectively. Other frequently examined factors included previous violence (nine studies\textsuperscript{55, 66-68, 70-72, 77, 79}, 44\% positive finding), young age at time of offence (eight studies\textsuperscript{3, 52, 65, 68, 69, 71, 72, 74}, 50\% positive finding), employment (eight studies\textsuperscript{54, 68, 70-72, 74, 75, 79}, 34\% found that it was negatively associated with reoffending, the remainder finding no association), previous psychiatric
admissions (10 studies\textsuperscript{62, 67, 69-76}, 10\% found positive effect) and substance use (seven studies\textsuperscript{3, 62, 65, 71, 75, 76, 80}, 43\% positive finding).

In terms of diagnostic groups, personality disorder (PD) was examined by nine studies\textsuperscript{3, 68-72, 73, 80, 81}, with 78\% of studies finding a positive association with reoffending. Six studies examined psychosis\textsuperscript{84, 62, 68, 71, 74, 75}, with 50\% finding this was negatively associated, and the remainder finding no association, with reoffending. However, four studies\textsuperscript{82, 55, 66, 67} found that “diagnosis” as a predictor (encapsulating PD and psychosis) was unrelated to reoffending, somewhat weakening these initially strong findings. This differential pattern of results likely reflects the diagnostic homogeneity of these four studies, in which the vast majority of patients had psychotic disorders and only small numbers were diagnosed with personality disorder (8\%, 8\%, 13\%, and 9\%, respectively), whereas studies which had more variance in diagnostic group, and thus more power to detect significant differences, tended to find positive results. For example, in a sample in which the number of participants with PD or psychosis was approximately equivalent\textsuperscript{81}, PD emerged as a factor associated with reoffending.

In addition, one study\textsuperscript{82} which examined ‘success of transfer’ from high security to medium security found no significant demographic predictors. This outcome was deemed conceptually distinct from any of the three main outcome groups (as an unsuccessful transfer could be due to inpatient violence or worsening of symptoms, for example) and thus the predictors were not included in the variable count. A list of the variables examined in this study is included in Supplementary Table 1.

\textbf{****** Table 2 and Figure 4 about here******}

\textit{b) Neuropsychological and Neurophysiological Predictors}

Six demographic studies examined the effect of IQ on reoffending\textsuperscript{70-72, 74, 75, 79}. It is notable, however, that these studies did not conduct a formal assessment of IQ; scores were extracted from patient files which may have limited the findings in terms of standardising the assessment tool used. Further variation may also have been introduced in terms of when the assessment was conducted (i.e. at admission, during an acute phase of illness, during court proceedings, etc.), which was not evident from the reviewed papers. Five of these investigations found no relation to reoffending\textsuperscript{70-72, 74, 75}, while one study found a positive association (i.e. those with lower IQ were more likely to reoffend)\textsuperscript{79}. 
Howard and Lumsden\textsuperscript{83} assessed the relationship between the contingent negative variation (CNV) event related potential during a Go/No-Go task, and community reoffending in a sample of 44 admissions to a high-secure forensic hospital. The CNV during this task has been correlated with measures of impulsivity\textsuperscript{84} and has been used as evidence of pathological impulsivity in court proceedings\textsuperscript{85}. Thus, it can be considered an objective measure of behavioural impulsivity. Based on the CNV results obtained, patients were classified as high or low risk, dependent on whether their score was one standard deviation outside or within a control group’s score, respectively. At fifteen years post-testing, criminal records were examined to reveal that six of 21 in the high risk group had been convicted of another offence, including manslaughter, burglary and arson. This compares with only one of 23 in the low risk group, convicted of theft. Thus, it appeared that using the CNV during Go/No-Go was sensitive to differentiating those who may reoffend, and appeared to identify those at risk of committing more serious offences. The authors assert that the overall predictive accuracy was 63.6% and the relative improvement over chance was 72%.

c) Biological Predictors

No studies examining the effect of biological variables on community reoffending were identified.

Discussion

This systematic review of objective factors relating to outcomes in forensic mental health services is, to our knowledge, the first review of such factors to be conducted.

In terms of demographic factors, the predictors of inpatient violence included previous psychiatric admissions (67% positive finding), with mixed findings for young age (50% found an association with inpatient violence). Demographic factors associated with an increased length of stay included the severity of the index offence (90% positive finding) and having a history of absconding (67% positive finding). Initially psychosis appeared to be associated with an increased length of stay, however once studies examining ‘diagnosis’ as a predictor more broadly were considered, this association was weakened, probably due to sample diagnostic homogeneity as a low number patients included in these studies were not diagnosed with psychosis. Our findings relating to reoffending suggest previous offending, young age at admission or discharge, and personality disorder are relatively robust predictors of recidivism with the large majority of studies examining each factor indicating a positive association. The majority of studies examining psychosis found that this had no relationship with future offending, perhaps reflecting the relative efficacy of treatments that are available for psychotic disorders in comparison to personality disorder.
This review may have been limited in its ability to examine demographic predictors of outcome, as it excluded papers relating to risk assessment tools, which focus on this type of predictor. Structured professional judgement tools such as the HCR-20\(^6\) include items such as young age, identified by this review to be related to future offending, suggesting that they do hold useful predictive properties. However, many factors identified in this review showed conflicting results, for example young age was found to be associated, and not associated, with inpatient violence in an equal number of studies, just as a previous prison sentence was found to increase the length of stay in two studies, but found to be unrelated in two further studies. This suggests that demographic factors in isolation are not particularly useful to clinicians in assisting decision making, but may perhaps hold more validity when considered in combination (as risk assessment tools advocate).

In addition, demographic factors are static and thus not sensitive to changing risk which may be picked up by indices of neurological or biological function. A further limitation relating to the demographic results is that combining similar, but perhaps slightly different demographic factors (e.g. ‘severity of offence’ and ‘a violent or homicide offence’), may have somewhat distorted the true relationship between a given predictor and outcome. Future research should aim to operationalise predictor variable definitions to aid in the understanding of the unique contributions each predictor makes. This criticism also holds in relation to the definitions of outcome. For example, inpatient violence often has broad and differing conceptualisations in research investigations\(^86\), and although the majority of papers included in this review included episodes of both verbal and physical aggression in this outcome category, some excluded verbal threats\(^27\), and some included specific operationalisations such as “throwing food or an object that strikes another person”\(^23\). Length of stay may also have different implications across countries. For example, in the UK length of stay is linked to clinical responsiveness. Patients admitted under a hospital order are able to move from hospital to conditions of lesser security once they are deemed to have responded to treatment and reduced their level of risk. However, this may not be the case in other countries such as the USA where fixed length sentences may have been imposed. In this review one third of studies examining length of stay were conducted in the USA, with 50% conducted in Europe and 17% in Australasia. To allow greater insights into our findings, information about the location of individual studies has been included in Supplementary Table 1.

Common themes emerged from the identified neuropsychological and neurophysiological predictors; impulsivity as assessed by the contingent negative variation event related potential was associated with future reoffending upon discharge\(^83\), and SCWT errors (poor cognitive inhibition)\(^32\) were associated with increased frequency and severity of inpatient violence. Both of these facets could be considered to reflect poor behavioural controls, and thus this may
be an area which merits further research in relation to its utility as a marker of violence or reoffending. One study included in this review did not support this assertion, however as previously discussed it was underpowered, with a very short follow-up period and a low rate of inpatient violence was observed. Poor social cognition emerged from two studies as a robust marker of outcome. Misperception of angry voices was found to be associated with inpatient violence, and another study identified poor reading of emotion from the revised eyes task to be the overwhelmingly best predictor of risk and unmet need at follow-up. These results indicate that both cognitive and social-cognitive deficits appear to be associated with outcome, and could be targets for effective treatment.

The strategy of using neuropsychological tests to predict outcome is strengthened by other, non-prospective, studies not included in this review. For example, it was shown that scores from the Iowa Gambling Task could be used effectively to predict whether MDOs had been secluded in the past for either predatory or impulsive violent acts while in secure mental health services. However, one cross-sectional study found no significant association between neuropsychological measures and previous inpatient violence in 82 violent men with schizophrenia (including the National Adult Reading Test, the Wechsler Abbreviated Scale of Intelligence, Stop Task and the CANTAB-2 battery), although current and predicted IQ tended to correlate negatively with the number of violent incidents across an individual’s time in hospital, suggesting that there may be a role for neuropsychological function in the emergence of violent behaviour. More prospective studies are required to fully elucidate relationships such as these.

The use of biological markers to assist in clinical decision making also appears to have support from the reviewed studies. Both serum cholesterol and creatine kinase appeared able to predict inpatient violence to a reasonable degree of accuracy. Low serum cholesterol has been linked to higher rates of death from violence or suicide, and experimentally lowered cholesterol has been linked to aggressive behaviour in animals. A putative mechanism of action suggests that low cholesterol reduces the integrity of cell membranes, making serotonin receptors less efficacious, and poor serotonergic transmission has been linked to violent behaviour. Serum cholesterol as a marker has shown great promise in another prospective study of non-forensic inpatients; total cholesterol had a significant negative relationship with inpatient suicidal and violent behaviour, and to 3-month post-discharge violent behaviour. This is an area for future research and development with strong potential.

A number of other studies, not included in this review due to the samples being referred for ‘forensic psychiatric evaluation’ as opposed to admitted to services, have also investigated biological markers and show some promise.
For example, one study\(^7\) found that 27% of variance in reoffending could be explained by low non-oxidative glucose metabolism in a sample of violent offenders referred for evaluation and followed up eight years later. Another study\(^8\) showed that high levels of the thyroxine hormone triiodothyronine was associated with relapse into offending in another cohort of offenders referred for psychiatric examination. The use of these markers to predict other outcomes such as inpatient violence in individuals specifically detained in forensic mental health services is an area to be explored further.

The use of biomarkers to predict complex behavioural outcomes such as aggression or reoffending requires ethical consideration. Biomarkers in psychiatry have been subject to ethical scrutiny, namely for reasons including oversimplification of multifaceted and complex conditions, and by shifting the focus of ‘risk’ to the individual as opposed to considering the wider societal contributing factors\(^9\). These issues are relevant to MDOs, and further work in this area should be mindful of the wider implications of the findings. Certainly at this early stage, putative biomarker predictors should be considered alongside clinical judgement and other predisposing factors such as personality pathology. There are also scientific issues to be resolved before the use of biological markers can be condoned. For example, an acceptable level of sensitivity and specificity would need to be established for any putative marker, and this would need to add incremental validity to any risk assessments that are currently in practice. An idea of the temporal stability would also be required, i.e. over what time frame does this marker suggest a risk? Interactions with medications and the ‘trait’ vs. ‘state’ status of any biomarker would be further considerations before widespread use could be advocated.

In conclusion, the findings of this review suggest that using neuropsychological, neurophysiological and biological markers to inform outcome is a feasible and potentially useful strategy. However, development of such markers is in its infancy and further research in this field is required to translate these findings to clinical practice. Initial replication of the promising, small scale studies identified in this review are needed and, if successful, large prospective cohort studies would be essential to establish the merit of such a strategy. Once developed, adding empirical markers such as these to clinical decision making tools may be a beneficial strategy in the future to improve outcomes for MDOs, who are a group at present experiencing lengthy admissions to psychiatric care and poor outcomes in terms of reoffending. Thus, innovation in this area is essential.
Acknowledgements

Ottilie Sedgwick receives funding support from the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre at South London and Maudsley National Health Service (NHS) Foundation Trust and King’s College London. Veena Kumari is part funded by the Biomedical Research Centre for Mental Health at the Institute of Psychiatry, Psychology and Neuroscience King’s College London, and the South London and Maudsley NHS Foundation Trust, UK. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. The authors wish to thank Hannah Mullens for reading and providing constructive comments on an early draft of the manuscript.
References

25. Dietz PE, Rada RT. Battery Incidents and Batterers in a Maximum Security Hospital. Arch Gen Psychiatry 1982;39:31-34.
64. Buchanan A, Leese M. Quantifying the contributions of three types of information to the prediction of criminal conviction using the receiver operating characteristic. Br J Psychiatry 2006 May;188:472-478.
22.

68.

69.

70.

71.

72.

73.

74.

75.

76.

77.

78.

79.

80.

81.

82.

83.

84.

85.

86.

87.

88.

89.


