Author’s response to reviews

Title: Prospective study of factors influencing conditional discharge from a forensic hospital: the DUNDRUM-3 programme completion and DUNDRUM-4 recovery structured professional judgement instruments and risk.

Authors:

Mary Davoren (davorenm@tcd.ie)
Zareena Abidin (zareena.abidin@hse.ie)
Leena Naughton (leenanaughton@gmail.com)
Olivia Gibbons (olivia.gibbons@yahoo.com)
Andrea Nulty (andrea.nulty@hse.ie)
Brenda Wright (brenda.wright1@hse.ie)
Harry G Kennedy (kennedh@tcd.ie)

Version: 2 Date: 6 January 2013

Author’s response to reviews: see over
Dear Editor

My colleagues and I have used the delay since first submitting this article to lengthen the follow-up period. We believe this has considerably strengthened the statistical power of the study and also assists us when addressing some of the helpful points raised by the referees. We have carefully addressed each point raised by each of the reviewers. We hope the revised manuscript is now acceptable for publication. We believe it has benefitted from the constructive criticisms received and we wish to express our appreciation to the reviewers for the trouble they have taken.

We are of course happy to make any further modifications or address any further questions that might arise.

Yours sincerely

Harry G Kennedy

Professor Harry Kennedy
Executive Clinical Director
National Forensic Mental Health Service
Central Mental Hospital, Dundrum, Dublin 14, Ireland
And
Clinical Professor of Forensic Psychiatry
Department of Psychiatry
Trinity College Dublin

Phone +35312989266
Caroline Higgins, PA +35312157458
Fax +35312157579
Please copy e-mails to Caroline.Higgins2@hse.ie
This paper addresses an important area of forensic mental health. It exploits a natural experiment to investigate whether 2 new tools can predict the decision making of a newly constructed review board. Particularly, it is a test of whether the new tools, the DUNDRUM 3 and the DUNDRUM 4 can predict the behaviour of the review board. It is presumed that if so it may provide a basis to organize and construct recommendations to such boards in making release decisions. It compares the new tools with a range of other structured professional judgement tools.

The study is not blinded. Extensive results are presented using AUC as a measure of the various tools and their elements ability to predict who given a conditional discharge. The numbers are however, small. There were 56 eligible participants, of whom 7 were discharged. The study therefore is of the tools’ ability to predict these 7 release decisions from the other 49 rejections of release.

Response 1: we are grateful for the reviewer’s observation regarding numbers. Because we have now been able to increase the follow-up period, there have been 12 discharges and we believe this improves the reliability of the findings.

Major Compulsory Revisions

The study slightly misleadingly suggests that this is a study of 98 persons.

However only 56 were eligible for release, so reference to the larger number is
somewhat misleading and should be removed.

Response 2: we have now edited the paper and abstract to make it clear that 56 patients were eligible for release by the mental health review board e.g.

Page 8: Participants: All patients in the hospital were assessed between February and March 2011 of whom 56 were newly eligible for conditional discharge to the community by the mental health review board.

Further, of the 56, 8 were unfit to stand trial. It is not clear [on page 10, see also page 15 and page 18 where they state the new legislation only covers those NGRI] whether the issues considered by the Board for those persons unfit to stand trial are the same as for the other 48 NGRI persons. Thus should this really be a study of just 48 persons? The authors need to clarify this point. Responses to this issue is essential as this is a small study of a natural experiment.

Response 3: we have now clarified this as follows –

Pages 8-9: All patients in the hospital patients were assessed between February and March 2011 of whom 56 were newly eligible for conditional discharge to the community by the mental health review board. Only those detained under the Criminal Law (Insanity) Act and eligible for review by the MHRB were included (figure 1). This excluded 16 patients detained under the Mental Health Act 2001 who were reviewed by a different body, the Mental Health Tribunal, which did not have a power to grant a conditional discharge or similar order. Data on the patients detained under the Criminal Law (Insanity) Acts 2006 & 2010 and their subsequent discharges from the hospital were then gathered up to 31st December 2012 (Fig. 1). Twenty three detained under other sections of the Criminal Law (Insanity) Act who were remanded or sentenced prisoners transferred from prison to hospital were not eligible for absolute or conditional discharge to the community by the Mental Health Review Board. There were therefore 56 patients (8 detained as unfit to stand trial (UST), 48 not guilty by reason of insanity (NGRI)) who were eligible for discharge or conditional discharge by the Mental Health Review Board. Those detained because UST were eligible for conditional discharge by the MHRB in the same way as those detained because found NGRI, though those detained because UST could also be returned to court by the treating consultant if they regained fitness.

And

Page 13: Of the 8 patients admitted because they were deemed unfit to stand trial 5 remained as in-patients during the period of observation while three were found fit to stand trial in court and returned to prison or the community. All had been considered by the Mental Health Review Board but had not been discharged. Of the 48 patients who had the legal status of Not Guilty by Reason of Insanity (NGRI) 12 were discharged subject to conditions by the MHRB.
while the remaining 36 continued to be detained as in-patients in hospital at the end of the period of observation (fig 1).

The authors further acknowledge that it may be atypical as the release of some of these persons was delayed by the new legislation, and thus may not be typical in the future of persons being considered for release.

Response 4: we agree and made this point in the ‘limitations’ section. We believe the longer period of follow-up has to some extent overcome this problem. We have now added this comment to the relevant paragraph under ‘limitations’ on page 20 –

Because these patients had been delayed so long, many of the patients achieving conditional discharge had near perfect scores on the DUNDRUM-3 and DUNDRUM-4 scales. We believe this problem has to an extent been overcome by the period of follow-up, examining conditional discharges after three MHRB hearings rather than just the first hearing under new legislation. In the future patients may be offered a conditional discharge with less perfect scores.

A further indication that matters had ‘settled down’ may now be obvious only to the reviewers who have had the advantage of seeing the previous draft. The first six conditionally discharged tended to be older and to have had longer admissions prior to discharge. The twelve now reported had negligible differences in age and length of stay from those not discharged.

Minor Essential Revisions

There are some minor typographical errors. Further, the authors need to be a little tighter with their use of language of what they studying: the ability to predict review board decisions, and the factors that appear to predict those decisions.

They at times suggest it is a study of clinicians, which it is not.

Response 5: We acknowledge the importance of this point. We have carefully edited to ensure that we have been consistent in distinguishing between the ‘recommendations’ of psychiatrists and the ‘decisions’ of MHRBs. Under ‘methods’ on page 7 we have described the process –

Structured professional judgement instruments and measures of mental state and global function were assessed at baseline in February 2011 just before the commencement of the Criminal Law (Insanity) Act 2010 which permitted conditional discharge for the first time. The Mental Health Review Board was supplied with a report from the treating consultant psychiatrist, had access to all clinical notes, assessments and reports and heard oral evidence from the patient, the treating consultant psychiatrist and the patient’s lawyer, and any other evidence the MHRB wished to hear. Neither the clinicians nor the members of the MHRB could be blinded to the information used to make the ratings for the various instruments, nor could they have been blinded to such information. Clinicians decided to recommend or not to recommend absolute or conditional discharge in accordance with their normal practice, a synthesis of the results of structured professional judgement instruments such as the HCR-20
and broad bio-psycho-social assessment including reports of progress in hospital. The MHRB received reports from the treating psychiatrists and other clinicians and was independent in the exercise of its powers [15].

And on page 22:

This is a study of the overall process by which clinicians make recommendations and MHRB members decide who to discharge. We have shown that clinicians and decision makers appear to take into account more than risk alone when deciding to release a patient from a secure hospital setting.

Discretionary Revisions

The authors present a large amount of data about the ability of the tools, which are strongly inter-correlated, to predict conditional release. This has value given the preliminary nature of this work, but either being clear about which ones are the greatest value or have the largest contribution would be of value.

Response 6: we agree that this is a very important point. We have included a logistic binary regression analysis, which to some extent cast light on this important subject –

Pages 15-16: Logistic Regression

Binary logistic regression was performed using forward stepwise likelihood ratios for the variables with significant areas under the curve. These were the DUNDRUM-3, DUMDRUM-4, HCR-20 total score, S-RAMM current score, START-S, SAPROF, GAF and PANSS total scores. The START-V was omitted and HCR-20 scales combined to reduce multiple co-linearity while the START-S was also omitted to avoid co-linearity with SAPROF. The iterative process resolved in one step for the GAF as the only significant variable remaining in the regression equation. The model was acceptable with Hosmer and Lemeshow $X^2=8.45$, df=8, $p=0.391$, Nagelkerke $R^2=0.610$, and GAF odds ratio(OR)=1.258 (95% confidence interval 1.089 – 1.454) $p=0.002$. This model correctly predicted 95.1% of those not discharged, 83.1% of those who were discharged and 92.5% overall. This model was not robust with backwards logistic regression yielding a different result, resolving in five steps (DUNDRUM-3 odds ratio=0.682, 95% CI 0.470 – 0.991, $p=0.045$; SAPROF OR = 0.633, 95% CI 0.413 – 0.967, $p=0.034$; GAF OR = 1.387, 95% CI 1.054 – 1.825, $p=0.020$, Hosmer and Lemeshow $X^2=6.05$, df=8, $p=0.889$, Nagelkerke $R^2=0.739$, correct predictions as before).

Because the model was not robust, the logistic regression was repeated, this time omitting the GAF. This was done because of the apparent interactive effects between the GAF and SAPROF, evident also when the START-S was included instead of the SAPROF. Backward logistic regression resolved in five iterations, Nagelkerke $R^2=0.603$, Hosmer and Lemeshow $X^2=6.05$, df=8, $p=0.642$, with 97.6% correct prediction of those not discharged, 66.7% correct prediction of those discharged and 90.6% correct overall. The model included two variables DUNDRUM-3 OR = 0.717, 95% CI 0.561 – 0.916, $p=0.008$ and PANSS total score which did not reach significance OR = 0.913, 95% CI 0.821 – 1.016, $p=0.097$. This model
was reasonably robust and when repeated by forward logistic regression yielded a model in one iteration with Nagelkerke $R^2 = 0.525$, Hosmer and Lemeshow $X^2 = 2.37$, $df=7$, $p=0.927$, 92% correct prediction of those not discharged, 66.7% correct prediction of those discharged and 86.8% correct overall and included a single variable, DUNDRUM-3 odds ratio = 0.698, 95% CI 0.563 – 0.866, $p<0.001$.

We have further discussed this on page 19

It is particularly interesting that the Global Assessment of Function (GAF) emerged as one of the strongest predictors of conditional discharge, and dominated the logistic regression models. The DUNDRUM-3 and DUNDRUM-4 correlated strongly with the GAF (Spearman $r$ for DUNDRUM-3 and GAF, $r = -0.804$, $p<0.001$, DUNDRUM-4 and GAF $r = -0.729$, $p<0.001$) and the DUNDRUM-3 programme completion scale emerged as the next strongest predictor. It is possible that the items of these scales represent elements of what is taken into account when rating the GAF, or it may be that an element in the rating of all of these items, like the GAF, is a rating of some ‘structured intuition’ as recently suggested [28]. The ratings of the seven items of the DUNDRUM-3 programme completion scale and the six items of the DUNDRUM-4 recovery scale have in common that they are based on motivation theory, cycle of change and engagement.

It is important to emphasise that this is only predicting the community release decision, not successful community tenure or re-offending.

Response 7: we agree, and have emphasised this on page 19 as follows

**Limitations**

*This was a study of factors influencing the decision by a mental health review board (MHRB) to conditionally discharge, it was not a study of adverse outcomes following discharge.*

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I declare that I have no competing interests
Reviewer's report

Title: Prospective study of factors influencing conditional discharge from a forensic hospital: the DUNDRUM-3 programme completion and DUNDRUM-4 recovery structured professional judgement instruments and risk.

Version: 1 Date: 31 October 2012

Reviewer: Philip Sugarman

Reviewer's report:

The question being address is not well defined - is it to predict discharge, or to report on the properties of these tools.

Response 8: This is an interesting point for clarification. We have now clarified this in the statement of hypothesis at the end of the introduction –

Page 6: In this prospective cohort study we examined whether the DUNDRUM-3 programme completion and DUNDRUM-4 recovery instruments [10] measured at baseline just before the commencement of the new legislation could distinguish between those who were or were not subsequently discharged either conditionally or unconditionally from this high and medium secure forensic hospital setting, and whether measures of risk, need for therapeutic security, mental state and global function also played a part in the decision.

We have also extensively referenced the earlier papers in which the properties of the instruments have already been reported. However the reviewer is correct that we have given a lot of space to the description of ‘item to outcome’ results, because we believe this is more helpful in a study concerning structured professional judgement than the reporting of results for overall scores, since the use of overall scores would merely be actuarial. We have now clarified this intention in the hypothesis section as follows –

Page 6 contd: Because we are interested in the use of structured professional judgement itself, we also wished to study the ‘item to outcome’ results for these instruments, since a study of the scale scores would be insufficient, amounting to an actuarial study only.

The methods are appropriate and well described, and the data is sound.

Standards for reporting and data deposition are good. The statistical techniques are used appropriately.

The discussion and conclusions are little light on what the study is telling us about the predictability and predictors of discharge, and about clinician and tribunal behaviour.
Response 9: see Response 7 above regarding the use of binary logistic regression. We hope this addresses the point about predictability and predictors.

Response 10: we share the reviewer’s interest in the distinction between clinician and tribunal behaviour. We believe this would make an interesting study for the future, though obtaining the consent of tribunals to study and report on them might be problematical. We have however acknowledged the point under ‘limitations’ as follows –

Page 20: This study did not attempt to distinguish between the recommendations of the treating psychiatrists and the decisions of the mental health review board (MHRB) since that would have required greater numbers and access to the deliberations of a judicial body.

The relevant literature could be more widely cited in what is a significant study.

Response 11: we have now cited the recent paper by Carroll (reference 29) which we feel casts an interesting light on the results of the logistic regression.

Page 19: It is particularly interesting that the Global Assessment of Function (GAF) emerged as one of the strongest predictors of conditional discharge, and dominated the logistic regression models. The DUNDRUM-3 and DUNDRUM-4 correlated strongly with the GAF (Spearman r for DUNDRUM-3 and GAF, r = -0.804, p<0.001, DUNDRUM-4 and GAF r = -0.729, p<0.001) and the DUNDRUM-3 programme completion scale emerged as the next strongest predictor. It is possible that the items of these scales represent elements of what is taken into account when rating the GAF, or it may be that an element in the rating of all of these items, like the GAF, is a rating of some ‘structured intuition’ as recently suggested [28]. The ratings of the seven items of the DUNDRUM-3 programme completion scale and the six items of the DUNDRUM-4 recovery scale have in common that they are based on motivation theory, cycle of change and engagement.

We could have expanded considerably on this theme but we feel we would merely be stealing Carroll’s thunder. We hope to return to this subject in a subsequent paper.

The limitations of the work are clearly stated, but could be more creatively addressed by focussing on what is surprising or otherwise about the predictors.

Response 12: we hope this has now been satisfied by the logistic regression and related discussion.

The future of the study for re-conviction should also be described a little more.

Response 13: under ‘generalisability’ we have commented

Page 22: The HCR-20 has been shown to predict those who have adverse outcomes on discharge from forensic hospitals [5,6,7]. Future research will be required to assess whether or not the DUNDRUM-3 programme completion and the DUNDRUM-4 recovery scales can predict those patients who will succeed in the community setting and those patients who will require to be recalled to the forensic hospital setting. It may be that different sets of items within the various structured professional instruments and symptom inventories will predict violent recidivism, relapse or readmission. A follow up study over a prolonged period will be
necessary to examine that question. We believe that a multi-centre, international study may be needed to achieve statistical power for a sufficiently detailed analysis.

We hope to return to this topic in a forthcoming editorial.

Acknowledgements appear to sufficient.

The title is OK, the abstract should be reviewed when the paper has been revised.

The standard of writing is good, if dry.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I declare that I have no competing interests
Reviewer’s report

Title: Prospective study of factors influencing conditional discharge from a forensic hospital: the DUNDRUM-3 programme completion and DUNDRUM-4 recovery structured professional judgement instruments and risk.

Version: 1 Date: 19 November 2012

Reviewer: Michael J Vitacco

Reviewer’s report:

It was my pleasure to review the paper entitled, “Prospective study of factors influencing conditional discharge from a forensic hospital: The DUNDRUM-3 programme completion and DUNDRUM-4 recovery structured professional judgment instruments and risk” for BMC Psychiatry. This is an interesting naturalistic study that came about due to a change in law that allowed extensive data to be collected on an understudied population. There are several strengths to this paper.

However, there is a major issue with the paper that requires a substantial change in the manner the paper is conceptualized. Notably, this is not a predictive study at all. This is a study where several risk instruments were utilized and the results shared with a committee that ultimately made discharge decisions. This factor is noted in the Discussion; however, the extent this poisons the study is minimized. Given the confound it is not surprising there were AUCs in the .90s.

Response 14: We are grateful to the reviewer for his comment. We can find only three AUCs of 0.9 or greater in all 9 tables, only one (the GAF, which we are certain could not have been seen by clinicians or MHRB) in table 10. We have now clarified the extent of this limitation as follows on page 29 –

**Limitations**

The ratings of the instruments and their component items were carried out by researchers who were blind to each others’ ratings and the research ratings were not provided to the clinicians or MHRBs. The success or failure of the instruments as statistical predictors depends on the extent to which they emulate the ways in which the clinicians and MHRBs make their recommendations and decisions. The information contained in the structured professional judgement instruments was inevitably available to the clinicians and MHRBs who could not be blinded to it. What was studied is the extent to which the instruments and
their contents predict the behaviour of the decision makers. The statistical inference is that the decision makers relied to a greater or lesser extent on this information. This was an observational study and therefore no special effort was made to blind clinicians who gave evidence or the members of the MHRB to such information. Since the purpose of the study was to identify those items that predict conditional discharge, it would be self-defeating and probably impossible to blind clinicians and MHRB members to such information.

We hope this satisfies the reviewer’s observation.

This paper would work better if it were simplified and thought of as a descriptive study of individuals considered for and granted conditional release. Otherwise, you are employing statistics designed for prediction in a study that has a fatal flaw.

Response 15: we have taken this comment on board and we have at every opportunity emphasised that this is a naturalistic descriptive study. It is also a prospective study in which the decision makers were blind to the research ratings. We acknowledge that this is a limitation inherent to naturalistic studies but not a ‘fatal’ flaw. Indeed, a naturalistic study of this important topic would not be otherwise possible.

There are other minor issues that include the following:

1. Not mentioning the Risk-Need-Responsivity factor for conditional release,

   Response 16: we have now mentioned this at an early stage in describing the process of conditional discharge –

   Page 7.....before the commencement of the Criminal Law (Insanity) Act 2010 which permitted conditional discharge for the first time. This allowed the MHRB to grant conditional discharge where appropriate and safe to do so, and subject to conditions regarding both treatment adherence and supports according to the individual’s risks and needs.

2. State the importance of risk assessment in the introduction.

   Response 17: this has now been covered, see below –

   Pages 4-5: Risk assessment is a key part of the process when making decisions regarding a patient’s readiness for discharge to the community. Dolan and Khawja [5] showed that the HCR-20 was a predictor of readmission and self reported violence in groups of discharged male medium secure patients. It has also been shown that the HCR-20 is a good predictor of both violent and non-violent offending in patients released from forensic psychiatric hospitals [6]. Doyle at al have also shown that dynamic (“current” and “risk”) items on the HCR-20 significantly improved the accuracy of prediction of violence after discharge from forensic units [7]. It has been shown that factors such as a higher score on PCL-R and a younger age at the time of first criminal offence were significantly related to release recommendations in a group of NGRI patients in the USA [8]. However evidence presented to Criminal Law (Mental Health) Review boards, when recommending a
patient for a move to less secure places, consists of more than risk assessment. In practice evidence given by clinicians also includes or takes account of factors such as rapport, insight, therapeutic alliance and use of leave from the hospital. Risk assessment in forensic psychiatry has evolved in recent years from the use of unstructured clinical judgement to the use of structured professional judgement instruments. These instruments add to the transparency and accountability of the process of risk assessment, uniting the tasks of prediction, assessment, clinical management and communication [9].

3. More complete description of the DUNDRUM measures.

Response 18: we have now added the following passage, which is similar to an equivalent passage in another of our papers -

The DUNDRUM-1 triage security scale is a static measure while the DUNDRU-3 and DUNDRUM-4 are dynamic measures sensitive to change and response to treatment. The content of these instruments is different from but complementary to risk assessment. When assessing readiness for discharge to the community, clinicians are likely to take more than risk assessment alone into account. Factors such as mental health, physical health, self care and activities of daily living, family and social networks, use of leave from the hospital and other such factors are all given strong consideration. These items are often included in clinician’s unstructured reports to mental health tribunals and review boards to assist these bodies in their decision making with regard to a patient’s readiness for conditional or absolute discharge, or neither. The items scored in the DUNDRUM-3 and DUNDRUM-4 include the above items and are based on motivation theory, cycle of change and engagement. The DUNDRUM-3 programme completion instrument consists of seven items - physical health, mental health, drugs and alcohol, problem behaviours, self care and activities of daily living, occupation, education and creativity and family and social networks. The six items of the DUNDRUM-4 recovery scale are stability, insight, rapport and working alliance, leave, HCR-20 dynamic risk items and victim sensitivity. Each item is accompanied by a series of definitions and rated from ‘0’ to ‘4’. A patient scoring mostly ‘4’s is unlikely to be ready for a move to a less secure place, a patient scoring mainly ‘3’s is likely to be ready to move from high to medium security, mainly ‘2’s is likely to be ready from medium security to PICU, mainly ‘1’s is ready for placement in the community setting and mainly ‘0’s is likely to be ready for an absolute discharge.

We have extensively referenced the handbook of the instrument itself and the web-site from which it can be freely downloaded (reference number 10). The headings for the content of each item are given in table 4. We have also extensively referenced earlier papers describing the content and psychometric properties. One of these (reference 13) actually included the content as additional material.

4. Adding age/ethnicity into the participants’ descriptions.

Response 19: we had already included the ages and lengths of stay. We have now mentioned ethnicity also -
At baseline, mean age was 43.7 years (SD 12.8) and mean length of stay was 11.1 years (SD 11.3) for the 56 eligible patients. All were Irish born.

5. Not sure the secondary analysis is needed unless you are going to make an argument that scales need to be changed based on item analyses.

Response 20: please see response 8 above, which we believe addresses the same point. We returned to this point towards the end of response 13. We hope this is satisfactory.

6. The paper could use a good editing (e.g., page 8, “Tio” should be “to”).

Response 21: we have attended to this.

7. The text contains a lot of extra information on the statistical analyses, which are also found in the tables.

Response 22: so far as possible we have now avoided unnecessary duplication. At times it is difficult to judge whether brevity or clarity are preferable. We are happy for the editor to sub-edit freely.

8. There are too many tables and that is hard to follow.

Response 23: we hope that response 8 above regarding the importance of item to outcome measures clarifies the reason for the number of tables. We returned to this subject also in response 13. We very much value the form of web-publishing pioneered by Bio-Med Central which allows authors to put all useful information before the interested reader.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Needs some language corrections before being published

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests' below.