Author's response to reviews

Title: The DUNDRUM-1 structured professional judgment for triage to appropriate levels of therapeutic security: retrospective-prospective validation study.

Authors:

Grainne Flynn (flynngr@tcd.ie)
Conor O’Neill (conor.oneill@hse.ie)
Clare McInerney (claremcinerney@ireland.com)
Harry G Kennedy (kennedh@tcd.ie)

Version: 2 Date: 4 October 2010

Author's response to reviews: see over
Dear Editor

My colleagues and I are grateful to the reviewers for their detailed critiques. We have accepted the points they raise and adjusted the text of the article accordingly. We will upload the revised text with 'tracked changes' in red. The notes below detail our responses to each reviewer point by point. Please note that as this was a continuous process over a number of days, some of the adjustments have been further altered in the light of comments by the other two reviewers, so the revised article is the full revision and may occasionally differ from the passages pasted here.

We have accepted the points they raise and adjusted the text of the article accordingly. We will upload the revised text with 'tracked changes' in red. The notes below detail our responses to each reviewer point by point. Please note that as this was a continuous process over a number of days, some of the adjustments have been further altered in the light of comments by the other two reviewers, so the revised article is the full revision and may occasionally differ from the passages pasted here.

We have taken the liberty of including extra figures illustrating the receiver operating characteristics. We hope the journal can accommodate these as they are a helpful visual illustration of the key statistical test described in the article. Should the editor prefer not to include them we would fully understand and the article would be complete without them.

We have now added a reference to the ethics committee approval and apologise for the oversight in omitting this from the earlier draft. We have also referenced the accompanying papers with the BMC editorial serial numbers.

My colleagues and I remain ready to respond to any further points the reviewers might wish to make.

Reviewer 1: Birgit Vollm

We appreciate the detailed comments of Dr Vollm. We have accepted and responded to each. For ease of reference we have pasted some of the text adjustments. 1. "It was not clear to me what the advantage of the new measure is over those identified as already developed from the literature" We have added brief commentaries on each of the existing instruments in the introduction and we have summarised the key advantages in the discussion. We hesitated to appear critical of other researchers, and we are perhaps still diffident (polite) about this.

"Eastman & Bellamy’s [8] Admission Criteria for Secure Services Schedule (ACSeSS)........... There are no published validation studies for these criteria and the assessment of 'likely duration' is unclear....p6

".....Other approaches have included Coid & Kahtan’s [10] algorithm .......which is specific to one jurisdiction; Shaw et al’s [11] structured professional judgment instrument ........using visual analogue ratings all rated using untethered Likert scales...; Sugarman & Walker’s [12] adaptation of the HONOS ........mixing patient centered items with institutional characteristics; and Collins & Davies’ [13] security centered factors such as physical security, relational security and procedural security with detailed item definitions, emphasising institutional characteristics over patient centered features. p7

and in the discussion -

"We believe this instrument has advantages over other instruments [8, 10, 11, 12, 13, 15] because it assesses patient centered rather than institutional factors, because, with a related
paper [22] it has been validated according to the criteria recommended by the Risk Management Authority of Scotland [23] and because it is drafted in a form that is likely to be applicable across jurisdictions and services. p21-22

2. We have now added a passage describing the development of the content of the questionnaire at pages 11-12

"The structured professional judgment instrument the DUNDRUM-1 is the product of an iterative drafting process. This commenced in early 2008 with a brainstorming and consultative session amongst the consultant forensic psychiatrists who are responsible for the decision to admit patients to the Central Mental Hospital. Nine consultant forensic psychiatrists were consulted, all of whom had worked or were working at the Central Mental Hospital. The nine had worked variously in nine medium or high secure forensic mental health services in five different jurisdictions. Colleagues in other disciplines were also consulted. The second phase consisted of an iterative process of refinement of definitions based on observation of discussions and practice at the weekly referrals meeting at the Central Mental Hospital at which all referrals are discussed and assessments prioritised. This meeting is chaired by the consultant forensic psychiatrist on call for that week and is attended by the leaders of all multi-disciplinary teams (consultant forensic psychiatrists), the heads of all disciplines (nursing, psychology, social work, occupational therapy), nurses in charge of wards and hospital managers. Clinicians from the psychiatric court liaison and prison in-reach service in the main remand prison also attend and those providing in-reach service in the main remand prison also attend and those providing in-reach clinics in the other prisons. Referrals for assessment with a view to admission from local mental health units are allocated to consultant forensic psychiatrists and when assessed these are also considered for admission at this meeting. The structured professional judgment instrument described here – the DUNDRUM-1 triage security instrument is part of the 21st revision of this draft. It forms part of a suite of structured professional judgment instruments [1] along with the DUNDRUM-2, an instrument for assessing the urgency of need for admission and prioritisation of waiting lists, and two instruments for assessing readiness for movement to less secure places, the DUNDRUM-3 programme completion instrument and the DUNDRUM-4 recovery instrument.

- and we have amended the reference to this in the abstract.

3. Does this mean all questions from the other identified questionnaires were included in this measure? Or were the themes identified and new questions formulated on these themes?

4. If questions were taken from the other identified questionnaires, it is essential that copyright issues are respected.....

5. Also the process of identifying "existing custom and practice" has to be explained in more detail.

We are grateful to Dr Vollm for the opportunity to clarify this. No content was taken from other questionnaires, only themes, many of which first appeared in our article of 2002. We were selective in the themes adopted and we also added themes, guided by the consultative phase described above. We have now substituted the word "themes" -
"We have collated all of these themes combined with a review of existing custom and practice to devise a manual [16]" (p7)

and we have deleted the word "existing" in the abstract. Dr Vollm may be raising an interesting point regarding the process of bringing the shared culture of expertise into a structured professional judgment instrument. We have discussed this at some length in the handbook. The handbook is referenced in the passage above [16] and is freely available via the web link. In the acknowledgments section at the beginning of the handbook we noted -

" This manual was written as a distillation of our training, experience and practice as forensic psychiatrists. Between us we have worked in four different countries so we hope that the structured professional judgement instruments contained here will work in a variety of health services and jurisdictions. With this in mind, the definitions, items and scales emphasise patient focused rather than institutional or local legal factors, in so far as possible. It is fashionable to say that a culture speaks through the authors of a text rather than the authors creating anything new. In this sense, any expertise we have drawn on is derived from a shared scientific culture, as described by Collins & Evans (2007). If this is the case then we hope that we are articulating a multi-disciplinary forensic mental health culture because many colleagues have contributed to this text through comments, criticisms and feedback, while many more have educated and acculturated us over the years, including many of our patients who are contributors to that culture."

We felt this was perhaps too general a point for this specific data paper, but we might return to it in a future review. We would of course we willing to add it to this article if the reviewer or editor wishes us to do so.

6. Statistics: there are a number of tests which appear to have been conducted but are not described in the methods section, including inter-rater reliability, internal consistency, factor analysis.

7. How were the cases included in the inter-rater reliability test selected.

We have now given greater detail as required, as follows -

"All data were entered for analysis in SPSS-16. All data was stored anonymously. p15

"Inter-rater reliability was calculated by arranging for a second rater (CO'N) to rate a consecutive series of cases (n=18) from a two week period, blind to the ratings of GF, who rated all cases. Cohen's kappa and Spearman's rank correlation test were calculated for items and the total scores were also correlated. p15

"Factor analysis was carried out using principle components analysis without rotation. Internal consistency was assessed using Cronbach's Alpha, examining whether the item to total correlations were improved by the omission of individual items. p16

8. Discussion: I felt the paper warranted a more in-depth discussion making more reference to other literature and alternative assessment instruments.
As outlined above, we have now been clearer in the Background and Discussion sections regarding the advantages of this instrument and its related suite of instruments over existing instruments or approaches. We have written a long introduction to the handbook. We are at some disadvantage as this is the first of a series of related articles validating this scale and three other scales. The discussion emphasises this. In another, related paper, we have discussed the ethical and organisational aspects of triage decision making and the advantages of the structured professional judgement approach. In this paper, we have emphasised the dynamic influence of resources on admission thresholds. In response to the reviewer's comments, we have now mentioned these topics in this paper.

9. "I did not understand how the authors arrived at the number of 316 for total assessment as there seem to be two samples of 100 and 246, totalling 346. Is this due to overlapping samples, i.e. those appearing in the 100 and the 246 were excluded...."

Yes. This is mentioned several times.

"In the two years January 2008 to December 2009 7,454 men newly committed to Cloverhill prison were screened by nurses at the point of reception using a four item screening questionnaire [18]. 1,454 were identified for full psychiatric assessment by the psychiatric in-reach and court liaison service [19]. Of these 100 were diverted from prison to psychiatric hospitals, including 27 sent to open wards in 16 local hospitals and 26 to low secure units (psychiatric intensive care or high dependency units) in three hospitals. These local hospitals are ‘approved’ to detain patients under the civil mental health act for Ireland, but are not ‘designated’ to detain patients under the criminal law insanity act for Ireland. A further 47 were diverted to the Central Mental Hospital, the sole forensic hospital for Ireland and the only hospital designated to receive patients detained under the Criminal Law (Insanity) Act 2006 for Ireland...." (p10)

and

"A related sub-sample was further analysed, consisting of 921 new committals (receptions) screened between April and June 2009, of whom 246 were identified for full assessment by the psychiatric in-reach team and 30 were diverted from the criminal justice system to any hospital placement. The 246 were rated using the DUNDRUM-1 security triage scale [16] by the same two clinicians in the same way. The 30 diverted from the criminal justice system to hospital overlap with the 100 described in the previous paragraph. A total of 316 were fully assessed and rated with the DUNDRUM-1. "(p10)

We have added the underlined sentence and some other clarifications to emphasise that these two samples overlap. The figure of 316 is arrived at by counting each individual once. Similarly we have clarified an earlier passage as follows -

"Study Design

This is a retrospective cohort study [17], instigated as part of the clinical audit and service evaluation process at the National Forensic Mental Health Service for Ireland. All those committed to a large remand prison (Cloverhill Prison, Dublin) were screened by nurses and a general practitioner and those identified as possible cases were referred for full psychiatric
assessment by a psychiatric prison in-reach and court diversion scheme. The clinical notes and assessments of all those assessed by the psychiatric prison in-reach and court diversion scheme over a three month period April to June 2009 were rated by two senior clinicians (GF and CO'N) blind to the eventual outcome. A further, overlapping sample identified in the same way (January 2008 to December 2009) consisted of all those who were diverted from the same remand prison to hospitals at various levels of therapeutic security over a two year period. " (p9)

10. Please describe briefly the relevant content of the "criminal law insanity act" .....  

We have now added as follows 

"A further 47 were diverted to the Central Mental Hospital, the sole forensic hospital for Ireland and the only hospital designated to receive patients detained under the Criminal Law (Insanity) Act 2006 for Ireland. This act permits the transfer from prison to the Central Mental Hospital of those remanded in custody or sentenced by the courts, if medically certified as having a mental disorder and in need of hospital treatment, as well as those found unfit to stand trial or not guilty by reason of insanity [19]."(p10)

This now includes a reference to a textbook combining the texts of the Acts in full with explanatory material.

Page 16, line 5: "X local" now replaced with the number, 16.

All remaining points regarding sub-editing have been corrected.

Missing data

There was no missing data. This reflects the relevance of the content of the items included. We have now commented on this -

"There was no missing data. This reflects the relevance of the item content" p14

"need to prevent access to contraband or victims" see also the definition in full in the additional material and in the handbook, as referenced at the end of this passage.

"Preadmission assessment should always be carried out by admitting services"

We agree that this is not always a legal requirement (as it is in England & Wales, though not in Ireland or Canada), nor is it always recognised as good practice. But our opinion is that a pre-admission assessment should always be carried out by the admitting service when using this instrument. Those jurisdictions that currently do not insist on a pre-admission assessment by the admitting service are generally recognised to be dysfunctional - but that is a debate for another place.

REVIEWER: Stuart Thomas
We are grateful to Dr Thomas for his detailed and constructive criticisms. We have accepted all the points and adjusted the text accordingly, as follows -

2. "the abstract was very vague and lacked detail. Specifically, the methods were confusing and the conclusions weak"

We have added detail in the methods and results section. We have now strengthened the conclusions section of the abstract in keeping with the discussion and conclusions section of the main article.

3. The background section is comprised of a series of statements and summaries but lacks any synthesis or critique and does not give any kind of rationale for this study or the development of the DUNDRUM-1

We have now strengthened the first paragraph -

"The clinical assessment of patients referred for admission to therapeutically secure and other hospitals has seldom been studied. The systematic allocation of patients to appropriate levels of therapeutic security is however central to the operation of mental health services generally and especially forensic mental health services. This is an area of clinical decision making that is critical for the timely delivery of services to those who are severely mentally ill and in prison or less secure hospitals. We set out to improve on unstructured professional judgement and existing instruments for assessment of need for therapeutic security. Our purpose is to provide a validated and reliable way of presenting such decisions in a transparent way. The structured professional judgement approach would also lend itself to benchmarking and quality standards.”

The remainder of the Background section has also been strengthened by pointing out the deficiencies of existing instruments (something we had been shy to do!).

4. What is a retrospective prospective study? Presumably this relates to the two different samples used here? This is by no means clear....

This is a synonym for a retrospective cohort study. We have now adopted the modern (American) nomenclature. We note that in an article recently published in BMC Psychiatry, the same design was simply described as a cohort study (Dolan M, Blattner R. *BMC Psychiatry* 2010, 10:76 (29 September 2010)).

We have now clarified the nature of the two samples/cohorts and their overlap, as outlined above.

5. How was the HCR-20 used for "cross validation" and why? Was it the same two clinicians who completed the DUNDRUM-1? Why do you think it was not correlated with DUNDRUM-1 scores?

We have added an explanatory sentence in the 'Background' section

"We hypothesised also that the DUNDRUM-1 triage security scale would correlate weakly or not at all with the HCR-20, a measure of risk which is not designed to take account of the seriousness of the risk or complexity of treatment need."
The clinicians who performed the assessments of DUNDRUM-1 and HCR-20 were the same, and were therefore blind to outcomes. The point raised regarding need for independence in assessment is addressed in much greater detail in an accompanying paper on cross-validation in in-patients. The point made in this paper is that the HCR-20 measures something different from need for therapeutic security. We have now added a discussion of this point -

"The scale does not correlate significantly with the HCR-20 ‘H’ score, a measure of static risk, nor does it correlate with the HCR-20 ‘C’ score, an assessment of dynamic risk factors. Further cross validation studies will be reported in a subsequent paper. This indicates that in measuring the need for therapeutic security, the DUNDRUM-1 measures something other than the risk of violence. Because the DUNDRUM-1 appears to have predictive validity, it would appear that it measures something considered in practice to be of greater importance than risk. A further cross validation study is reported elsewhere [22]. A prospective study to test the predictive validity of the HCR-20 in comparison with the DUNDRUM-1 would be needed to further clarify this." p22

6. There is no discussion about ethics i.e. accessing people’s case file information without their express consent.

We have now included a reference to the approval of the research and audit ethics committee.

"The study was approved by the local research ethics, audit and effectiveness committee"

p8

The committee was satisfied that since information was stored in anonymised form, and since the project formed part of an audit of effectiveness that was of direct benefit to service users, the impossibility of obtaining express consent for a retrospective cohort study should not prevent the project proceeding. Indeed it would be impossible ever to use such a design if express consent were required (see the discussion below about ‘real world’ research and utility). We note that other similar studies have been ethically approved elsewhere and published in BMC Psychiatry e.g. Dolan M, Blattner R BMC Psychiatry 2010, 10:76 (29 September 2010).

Dr Thomas’s comment raises another important ethical issue. We believe this study casts light on the very important ethics of clinical practice when making triage decisions and prioritising one patient over another. We have referred to this in the discussion and we have dealt with it in greater depth in a related paper on the DUNDRUM-2 triage urgency scale.

"We are conscious of the ethical and organisational aspects of triage admission decision making. We have discussed this in greater detail in a related article [23]." p22

7. The purpose of the statistical methods used is unclear and inconsistent with the aims / hypotheses, apart from perhaps predictive accuracy using AUCs……

The hypotheses are set out in the last paragraph of the Background section -

"Our hypotheses were that the eleven items taken as a scale would have acceptable psychometric properties and the total score would have predictive validity, distinguishing between those admitted to different levels of therapeutic security while each of the 11 items should also correlate with outcome. We hypothesised also that the DUNDRUM-1
We found that the DUNDRUM-1 has excellent psychometric properties. The DUNDRUM-1 has predictive validity for the level of therapeutic security to which mentally disordered remand prisoners were diverted. We have also demonstrated that each of the eleven items correlated with outcome. " p11

This is a topic that might be taken further in a future study using other forms of statistical analysis.

We hope this clarifies the "direction, statement and purpose". We have considered the reviewer's comment regarding trying to do too much in one paper. We hope it is now clear that we have stated our hypotheses and we have tested them. We feel it is of benefit to the reader to have this material, which is coherent, in one paper. We hope it is also easier for reviewers to understand the validation process in one piece. We believe this is one of the advantages of the BMC format over traditional paper-based journals in which numerous shorter papers each dealing with not very much, have to be published piecemeal in different journals over a period of years then gleaned by the reader from scattered data bases, many of which are not available freely on the web. Further parts of this validation exercise are in another paper, referenced in this paper. The two papers taken together fulfill the criteria for structured professional judgment instruments set out by the Risk Management Authority of Scotland and we have referenced this is a further passage (see below).

8. It is not clear how the authors came up with the 11 items or what the particular clinical utility of the DUNDRUM-1 actually is. Indeed, the authors seem to do a good job of negating any practical use of the tool in their discussion which talks about the wider service based practicalities of placement and how these kind of things work in the real world.
Please see the notes above and the additional paragraphs at p6 concerning the origins of the 11 items.

The particular utility of the DUNDRUM-1 is now set out in the Background section -

"The clinical assessment of patients referred for admission to therapeutically secure and other hospitals has seldom been studied. The systematic allocation of patients to appropriate levels of therapeutic security is however central to the operation of mental health services generally and especially forensic mental health services. This is an area of clinical decision making that is critical for the timely delivery of services to those who are severely mentally ill in prison or less secure hospitals. We set out to improve on unstructured professional judgement and existing instruments for assessment of need for therapeutic security. Our purpose is to provide a validated and reliable way of arriving at such decisions in a transparent way. The structured professional judgement approach would also lend itself to benchmarking and quality standards."

p5

and in the discussion section -

"The DUNDRUM-1 has predictive validity for the level of therapeutic security to which mentally disordered remand prisoners were diverted. We have also demonstrated that each of the eleven items correlated with outcome.....

"We believe this instrument has advantages over other instruments [8, 10, 11, 12, 13, 15] because it assesses patient centered rather than institutional factors, because, with a related paper [22] it has been validated according to the criteria recommended by the Risk Management Authority of Scotland [23] and because it is drafted in a form that is likely to be applicable across jurisdictions and services. ....

"This indicates that in measuring the need for therapeutic security, the DUNDRUM-1 measures something other than the risk of violence. Because the DUNDRUM-1 appears to have predictive validity, it would appear that it measures something considered in practice to be of greater importance than risk. A further cross validation study is reported elsewhere [22]. Though a prospective study to test the predictive validity of the HCR-20 in comparison with the DUNDRUM-1 would be needed to further clarify this..... (pp21-22)

Finally, we hope that our scientifically agnostic discussion of 'real world' confounding factors does not detract from the success of the instrument in this 'real world' test of predictive validity. We have added to the following paragraph in the Discussion -

"The data presented here reflects the dynamic interplay of population demand, the availability of alternative facilities in the community or at lower levels of therapeutic security [24] and resource allocation, so that actual admission thresholds arising from the balance of these are likely to vary from time to time and from one jurisdiction or administrative region to another. The area under the curve statistic is however a property of the instrument rather than a property of the population. The DUNDRUM-1 security triage assessment scale performs in practice as an actuarial predictor of the level of therapeutic security to which a person is allocated even in this 'real world' test of predictive validity. This is a naturalistic study of actual outcomes. Some of those admitted to PICUs may have had scores or clinical profiles more typical
of those admitted to medium security and vice versa (see tables 6 and 7) – this is the reality of clinical practice. Even under these constraints the scores were sufficiently distinct to predict the actual allocation to levels of security as measured by the receiver operating characteristic. Although the AUC is a property of the instrument, not of the population tested, the threshold scores are a property of the population studied and so the threshold scores found in this study might not generalize to another jurisdiction, for a number of reasons. The most important is that the threshold for admission to any given level of security will fall when more beds are available and will rise when the number of beds is smaller. A complex relationship can also be expected between the availability of beds at one level of therapeutic security and the demand for beds at adjacent levels [25] and these can influence time spent on waiting lists [26]. We believe that although the threshold may vary between services and jurisdictions, the threshold scores can be used to make valid benchmark comparisons between services and jurisdictions. " p23.

Finally, we believe this article is of interest to any reader involved in the teaching, development or use of structured professional judgment, anyone involved in the commissioning of forensic and general mental health services and anyone interested in waiting lists generally. It is of interest to health commissioners and the handbook has already been widely requested and downloaded e.g. by managers of the English High Security Hospitals.

REVIEWER: Gian Marco Polselli

We are very appreciative of this reviewer's assessment that the paper is of practical use. We appreciate the international perspective that is of prime importance to us in devising structured professional judgment instruments that should be useful across jurisdictions.

We note that the reviewer is familiar with the 'retrospective-prospective study' design.

1. The different samples...

We have now clarified the relationship between the two samples.

2. The researchers were blind to outcomes.

We agree with the reviewer that there is a need for an efficient structured instrument to evaluate the necessary therapeutic security level for such patients. We believe this paper describes the first predictive validity study of such an instrument.

We agree that aspects of the results might not generalise to other jurisdictions - e.g. the threshold scores, although the ROC/AUC statistic is a property of the instrument rather than the population and should be capable of replication in other jurisdictions.

We are grateful to the reviewer for his view that the manuscript's question was posed in a well defined way and that the methods were appropriate and well described. We hope the additional clarifications are helpful and do not detract from the original.