PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	DERIVATION AND VALIDATION OF A CLINICAL PREDICTION RULE FOR DELIRIUM IN PATIENTS ADMITTED TO A MEDICAL WARD: AN OBSERVATIONAL STUDY.
AUTHORS	Urreta, Iratxe ; Emparanza, Jose Ignacio; Martinez, Juan Antonio; Basabe, Iban; Belastegui, Ana; Aguirre, Cristina; Goikoetxea, Xabier; Liceaga, Nerea

VERSION 1 - REVIEW

REVIEWER	Camilla Wong, MD FRCPC MHSc Geriatrician, St. Michael's Hospital Assistant Professor, University of Toronto Associate Scientist, LKSKI of St. Michael's Hospital Canada I have no competing interests to declare.
REVIEW RETURNED	08-Jul-2012

REPORTING & ETHICS	Research ethics not stated. Was consent obtained from the study participants? If so, what was the recruitment rate from this consecutive cohort sample?
GENERAL COMMENTS	Overall, a useful paper in a much needed area. Introduction - This statement is made, "but its use has not become widespread in our setting since it requires variables that are difficult to measure on admission" -- Do we know this for certain? Any statistics to support this? Methods - It is unclear why the authors chose not to perform the CAM daily, but rather only upon consideration of discharge or after review of medical/nursing notes. - The CAM assessment was performed after "analysis of any relevant data in their medical record and nursing report"; if this is the case, how were the researchers potentially blinded to data such as patient age, functional status and medications which are usually information readily available from the medical record? - While a strength of the study was the use of a third assessor for the CAM in cases of disagreement, it would be useful to report the inter-rater reliability. - No comment on ethics approval. - No comment on recruitment rate (consecutive, but did all patients consent to the study?). Results

	- I believe the authors meant to use the word "various" rather than "variable" in the following sentence: "to those who had 2 or more points in the variable drugs" - In the text (rather than only the tables), it may be more useful to report the the cumulative incidence of delirium for point levels of 0, 1, 2, and 3 for the derivation and validation samples as the authors point out that "clinical prediction rules are useful tools for classifying patients at different levels of risk". - What was the result of the Hosmer Lemeshow test? This statistical method is described in the Methods section but the result is not reported. Discussion - This statement is probably more appropriate for the Results section: "Selecting the cut-off point for the highest sensitivity, the accuracy of the rule was: Se=93.4% (95% CI 85.5 to 97.2), Sp=60.6% (95% CI 54.1 to 66.8), PPV = 44.4% (95% CI 36.9 to 52.1), and a NPV = 96.5% (95% CI 92 to 98.5). The AUCs were 0.77 and 0.85 in the derivation and validation cohorts, respectively." - The word "be" is missing from this sentence: "It would interesting for the clinical predictive rule" - There is almost a two-fold discrepancy in the incidence of delirium between the derivation and validation cohorts. The authors should elaborate on reasons for this, such as a significant difference in dependency on ADLs. - The authors should comment on how their choice to not perform the CAM daily may have influenced the results.

REVIEWER

REVIEWER	Prof dr Bart Van Rompaey University of Antwerp - Belgium I declare to have no competing interests
REVIEW RETURNED	12-Jul-2012

- The reviewer completed the checklist but made no further comments.

VERSION 1 – AUTHOR RESPONSE

Introduction

We don't have any statistics to support the statement "but its use has not become widespread in our setting since it requires variables that are difficult to measure on admission", just the experience in our setting. In fact none of these previous rules is used in our hospital. When asking our colleagues why they don't use them, the answer is that the variables included in the rule are difficult to measure on admission.

Methods

With blinded to data we mean that they didn't know which potential predictors were selected for the study. Researchers who collected potential predictors were blinded to the outcome because the predictors were collected on admission.

The assessment of delirium is a component of the daily assessment of the patient by nurses. The researchers have used this nursing assessment in deciding whether the patient had suffered delirium or not on discharge.

We have added: The consent was obtained from the study participants and all patients gave their consent to participate in the study.

Results

We wanted to say drug variable or variable named drugs

We have added this: The patients with delirium of the two cohorts scored similarly: 17% and 7% scored 0, 48% and 30% scored ≤ 1 and 85% and 85% scored ≤ 2 in the derivation and validation cohort respectively.

We have added in table 7 The Hosmer–Lemeshow test (P = 0.873)

We have added: There is a significant difference in the ADL variables being those from the validation cohort more dependent then the derivation cohort. All the above mention variables explain the almost two fold discrepancy in the incidence of delirium between the two cohort.

The inter-rater reliability was not done.

VERSION 2 – REVIEW

REVIEWER	Camilla Wong, MD MHSc FRCPC Geriatrician, St. Michael's Hospital Assistant Professor, University of Toronto Associate Scientist, Li Ka Shing Knowledge Institute Canada I have no competing interests.
REVIEW RETURNED	26-Jul-2012

GENERAL COMMENTS	ABSTRACT - Participants -- I believe the authors are referring to "incidence of delirium" rather than "prevalence of delirium". Were patients admitted WITH delirium already at the time of admission also included in the study? METHODS - Validation cohort and model performance - This is not really a study of diagnostic accuracy, but rather a study of a prediction model. The authors should refrain from using "To assess the diagnostic accuracy of the rule ...". - It is unclear why the authors chose not to perform the CAM daily, but rather only upon consideration of discharge or after review of medical/nursing notes. - the CAM assessment was performed after "analysis of any relevant data in their medical record and nursing report"; if this is the case, how were the researchers potentially blinded to data such as patient age, functional status and medications which are usually information readily available from the medical record? - While a strength of the study was the use of a third assessor for the CAM in cases of disagreement, it would be useful to report the inter-rater reliability.

	RESULTS - Again, I believe the authors want to refer to "incidence" rather than "prevalence". DISCUSSION - Again, I believe the authors want to refer to "incidence" rather than "prevalence".

VERSION 2 – AUTHOR RESPONSE

Participants:
We have changed prevalence by incidence.
The patients admitted with delirium at the time of admission were not included in the study

Methods:
We have changed "to assess the diagnostic accuracy of the rule" by "to assess the predictive accuracy of the rule"
We chose not to perform the CAM daily for a mere operative question and in order to homogenize the criteria

You are right saying that the researchers had information about patient age, functional status and medications….but they didn't know which potential predictors were selected for the study.

We could not report the inter-rater reliability because we do not register it.

Results

We have changed prevalence by incidence

Discussion

We have changed prevalence by incidence