

## RESEARCH

# Development and validation of PRE-DELIRIC (PREdiction of DELIRium in ICu patients) delirium prediction model for intensive care patients: observational multicentre study

 OPEN ACCESS

M van den Boogaard *intensive care nurse, scientific researcher, PhD student*<sup>1</sup>, P Pickkers *professor in critical care*<sup>1</sup>, A J C Slooter *intensivist*<sup>2</sup>, M A Kuiper *intensivist*<sup>3</sup>, P E Spronk *intensivist*<sup>4</sup>, P H J van der Voort *intensivist*<sup>5</sup>, J G van der Hoeven *professor in critical care*<sup>1</sup>, R Donders *statistician*<sup>6</sup>, T van Achterberg *professor in nursing*<sup>7</sup>, L Schoonhoven *senior research fellow nursing science*<sup>7</sup>

<sup>1</sup>Department of Intensive Care Medicine, Radboud University Nijmegen Medical Centre, 6500HB Nijmegen, Netherlands; <sup>2</sup>Department of Intensive Care, University Medical Centre Utrecht, Utrecht, Netherlands; <sup>3</sup>Department of Intensive Care, Medical Centre Leeuwarden, Leeuwarden, Netherlands; <sup>4</sup>Department of Intensive Care Medicine, Gelre Hospitals, Location Lukas, Apeldoorn, Netherlands; <sup>5</sup>Department of Intensive Care, Onze Lieve Vrouwe Gasthuis, Amsterdam, Netherlands; <sup>6</sup>Department of Epidemiology, Biostatistics and HTA, Radboud University Nijmegen Medical Centre; <sup>7</sup>Scientific Institute for Quality of Healthcare, Radboud University Nijmegen Medical Centre

## Abstract

**Objectives** To develop and validate a delirium prediction model for adult intensive care patients and determine its additional value compared with prediction by caregivers.

**Design** Observational multicentre study.

**Setting** Five intensive care units in the Netherlands (two university hospitals and three university affiliated teaching hospitals).

**Participants** 3056 intensive care patients aged 18 years or over.

**Main outcome measure** Development of delirium (defined as at least one positive delirium screening) during patients' stay in intensive care.

**Results** The model was developed using 1613 consecutive intensive care patients in one hospital and temporally validated using 549 patients from the same hospital. For external validation, data were collected from 894 patients in four other hospitals. The prediction (PRE-DELIRIC) model contains 10 risk factors—age, APACHE-II score, admission group, coma, infection, metabolic acidosis, use of sedatives and morphine, urea concentration, and urgent admission. The model had an area under the receiver operating characteristics curve of 0.87 (95% confidence interval 0.85 to 0.89) and 0.86 after bootstrapping. Temporal validation and external validation resulted in areas under the curve of 0.89 (0.86 to 0.92) and 0.84 (0.82 to 0.87). The pooled area under the receiver operating characteristics curve (n=3056) was 0.85 (0.84 to 0.87). The area under the curve for nurses' and physicians' predictions (n=124) was significantly lower at 0.59 (0.49 to 0.70) for both.

**Conclusion** The PRE-DELIRIC model for intensive care patients consists of 10 risk factors that are readily available within 24 hours after intensive care admission and has a high predictive value. Clinical prediction by nurses and physicians performed significantly worse. The model allows for early prediction of delirium and initiation of preventive measures.

**Trial registration** Clinical trials NCT00604773 (development study) and NCT00961389 (validation study).

## Introduction

Delirium, characterised by an acute onset of fluctuating changes in mental status and changed levels of consciousness and inattentiveness,<sup>1</sup> has a high incidence rate in critically ill patients.<sup>2-4</sup> It is a serious disorder associated with prolonged stays in intensive care units and hospitals, higher costs, and increased morbidity and mortality.<sup>2 3 5</sup>

Several tools are available for assessing delirium in intensive care patients, of which the confusion assessment method—intensive care unit (CAM-ICU) has the highest sensitivity and specificity.<sup>6 7</sup> Screening intensive care patients is important,<sup>8-10</sup> so that timely treatment can be provided. However, preventive measures for delirium may also reduce its incidence, severity, and duration, as determined in other groups of patients.<sup>11 12</sup> General preventive measures in all intensive care patients are time consuming and may expose a substantial number of patients to unnecessary risks such as the adverse

Correspondence to: M van den Boogaard [M.vandenboogaard@ic.umcn.nl](mailto:M.vandenboogaard@ic.umcn.nl)

Extra material supplied by the author (see <http://www.bmj.com/content/344/bmj.e420?tab=related#webextra>)

effects of drug prophylaxis. Although several predictive models for non-intensive care patients exist,<sup>13,14</sup> as well as one for older medical intensive care patients,<sup>15</sup> no evidence based prediction model for general intensive care patients is available. The aim of our study was to develop and validate a delirium prediction model for intensive care patients and to determine its value compared with prediction by the attending nurses and physicians.

## Methods

### Study design

This was an observational multicentre study in which we firstly developed the PREdiction of DELIRium for Intensive Care patients (PRE-DELIRIC) model and then temporally validated it in a second prospective cohort in the same hospital. We then validated the model externally in four other Dutch hospitals.

### Development and temporal validation of the model

To develop the prediction model, we did a prospective cohort study in the Radboud University Nijmegen Medical Centre in the Netherlands. This study took place between 1 February 2008 and 1 February 2009. We did a second prospective cohort study in the same hospital for temporal validation of the model between 1 May and 1 September 2009.<sup>16</sup>

### External validation

After development and temporal validation, we externally validated the delirium prediction model with data from intensive care patients admitted to four other Dutch hospitals between 1 January and 1 September 2009. One of these hospitals was a university hospital, and three were university affiliated teaching hospitals; all had mixed intensive care populations (table 1). In these hospitals, trained intensive care nurses used the CAM-ICU at least twice daily.

### Delirium prediction by caregivers

To compare the predictive value of the model with that of the caregivers, we asked intensive care nurses and physicians caring for the patient to predict independently, within 24 hours after admission to intensive care, if patients would develop a delirious period during their complete stay in intensive care.

### Study population for development and validation studies

After the successful implementation of the validated Dutch version of the CAM-ICU,<sup>17</sup> the inter-rater reliability of the delirium screenings by the intensive care nurses was above 0.80 Cohen's  $\kappa$ , with a compliance rate of more than 90%, as described in more detail previously.<sup>10</sup> During the development and temporal validation studies, we included all adult patients admitted to the intensive care unit. To detect delirium, intensive care nurses screened all consecutive adult intensive care patients at least three times daily, and more often if required (for example, after sudden changes in behaviour, attention, or consciousness). This frequency of screening was in accordance with screening in daily practice. We excluded patients if they were delirious within 24 hours after admission to intensive care, had a sustained Richmond agitation sedation score (RASS) of  $-4/-5$  during the complete intensive care admission, stayed on the intensive care unit for less than one day, had serious auditory or visual disorders, were unable to understand Dutch, were severely mentally disabled, or had a serious receptive aphasia

or if the compliance rate of the delirium screening was less than 80% during a patient's stay in the intensive care unit.

To meet the same inclusion and exclusion criteria during the external validation study, we used consecutive patients with complete CAM-ICU screenings, defined as a CAM-ICU compliance rate above 80% per patient. We defined patients as having delirium when they had at least one positive CAM-ICU screening during their intensive care stay or were treated with haloperidol, as in these hospitals haloperidol is used only for treatment of delirium. To examine the predictive value of the PRE-DELIRIC model in daily practice in these hospitals, we did no compliance and inter-rater reliability measurements and used only data from CAM-ICU screenings as done in normal daily practice.

### Potential predictors

We collected demographic variables and information on potential risk factors identified by a recent systematic review.<sup>18</sup> We collected these data electronically within 24 hours of admission to intensive care (web appendix A). In addition, we included variables from the Dutch national intensive care evaluation database as potential risk factors when the delirium incidence rate associated with that variable was more than 50% higher than the incidence rate of the total group (web appendix B).<sup>19</sup> Wherever possible, the risk factors were collected as continuous variables (categorical or dichotomised when otherwise).

### Outcome definition

In view of our aim to develop and validate a delirium prediction model, the main outcome measure was development of delirium during patients' stay in the intensive care unit. We defined delirium as a minimum of one positive CAM-ICU screening during each patient's intensive care stay. In addition, we screened patients' medical and nursing files daily for signs of delirium.<sup>20</sup> If the files provided signs of delirium without a positive CAM-ICU screening or, conversely, if the files did not provide evidence of delirium and the patient had a positive CAM-ICU result, a delirium expert additionally screened patients according to the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (DSM-IV) criteria to rule out false negatives and false positives.<sup>1</sup>

### Data management and quality checks for development and temporal validation studies

We monitored the performance of CAM-ICU screening to ensure the quality of data collection. We calculated compliance as the percentage of assessments done each day in relation to the total number of assessments that should have been done. The mean compliance during the development and temporal validation studies was 90.4%. We measured the quality of CAM-ICU performance as the inter-rater reliability. For this, we compared the CAM-ICU screening assessed by the attending intensive care nurse with the CAM-ICU score assessed by an expert psychiatry nurse within a time window of one hour. We did 120 inter-rater reliability measurements, resulting in a Cohen's  $\kappa$  of 0.90 (95% confidence interval 0.82 to 0.98). The first author randomly double checked data from 15% of all patients included for completeness and accuracy.

### Statistical analysis

We calculated the sample size needed for the development of the model on the basis of the need for 10-15 delirious patients per risk factor plus 10% dropout. We imputed missing data for

the risk factors. Values were missing in the development study for urea (0.7%), liver enzymes (3.0%), bilirubin (18.0%), calcium (4.5%), sodium (0.3%), haematocrit (0.4%), metabolic acidosis (1.0%), and acute physiology and chronic health evaluation (APACHE)-II scores (0.7%). Data for all other variables were complete. All data for the temporal validation study were complete. We decided that if a laboratory measurement was not determined we had no reason to assume that the missing variable had an abnormal value, and we imputed the mean normal value. To calculate the normal value, we selected all patients with a normal value and then calculated the mean value for this group of patients and used it for imputation. When the APACHE-II score was missing, we imputed the mean value for the delirium or non-delirium group, depending on the results of the CAM-ICU. In the external validation dataset, 6.3% of the urea values were missing and imputed. For APACHE-II, 0.6% of the scores were missing, and we imputed a mean APACHE-II score for the group in the external validation set.

We used univariate logistic regression to develop the prediction model by assessing the association between each potential prognostic determinant and the presence or absence of delirium. We excluded determinants with a P value above 0.15 in univariate analysis or with a prevalence rate below 10%. With the remaining risk factors, we used multivariate logistic regression analysis with backward elimination (excluding risk factors with P values over 0.10) to evaluate the independent associations with the occurrence of delirium. The final model therefore contains independent risk factors for delirium. We estimated the prognostic ability of the model to discriminate between patients with and without delirium by using the area under the receiver operating characteristics curve (AUROC). We used bootstrapping techniques to adjust for overfitting—that is, for overly optimistic estimates of the regression coefficients of the risk factors in the final model. Two hundred random bootstrap samples resulted in shrunken regression coefficients of the risk factors and area under the curve of the developed model.<sup>21</sup>

In both validation studies, we multiplied shrunken regression coefficients for each risk factor by the observed patients' values. The outcome is a calculated predicted probability on which we built a new AUROC. Finally, to examine how well the model was calibrated, we calculated linear predictor values for each patient of every cohort by using the coefficients from the final development model. We used these linear predictors in a logistic regression model to test whether the prediction rule was well calibrated, resulting in a calibration slope and an intercept. A calibration slope of 1 and an intercept of 0 show a perfect calibration. Calibration plots for each cohort are available in web appendix D. We used SPSS 16.01, R statistics version 2.10.1,<sup>22</sup> using the rms package,<sup>23</sup> for all analyses.

## Results

### Development of prediction model

In total, we screened 2116 consecutive patients and excluded 503 of them (fig 1⇓). Of the remaining 1613 patients, 411 (25.5%) developed delirium. Table 2⇓ shows patients' characteristics, and web appendices B and C contain prevalence rates and delirium incidence rates for the separate risk factors. Of the 25 potential risk factors, we excluded alcohol misuse (7.8%), dementia (1.7%), use of an epidural catheter (2.2%), hyperamylasaemia (3.9%), hyponatraemia (5.8%), use of dopamine (0.2%), and use of lorazepam (0.7%) because of a prevalence rate below 10%. We excluded hypertension because of a P value above 0.15 in univariate logistic regression analysis.

After multivariate logistic regression analysis with the remaining risk factors, we constructed the PRE-DELIRIC model, which consisted of 10 risk factors (table 3⇓). The AUROC was 0.87 (95% confidence interval 0.85 to 0.89) and 0.86 after bootstrapping. Calibration of the model resulted in a calibration slope of 1.08 and an intercept of  $-0.06$ . The box shows the formula for the PRE-DELIRIC model.

### Temporal validation of prediction model

In the prospective validation study, we screened 748 consecutive patients and excluded 199 of them (fig 1⇓). Of the remaining 549 patients, 171 (31.1%) developed delirium (table 2⇓). The temporal validation resulted in an AUROC of 0.89 (0.86 to 0.92). The calibration slope of the temporal model was 1.2, and the intercept was 0.22.

### External validation of prediction model

We used data from 894 non-selected intensive care patients (table 2⇓) for external validation, resulting in an AUROC of 0.84 (0.82 to 0.87) with a calibration slope of 0.76 and an intercept of  $-0.59$ . The AUROCs of the four different hospitals did not differ from each other (data not shown). As no differences in prediction existed between the three studies, we pooled the data (n=3056), resulting in an AUROC of 0.85 (0.84 to 0.87) (fig 2⇓). The pooled data resulted in an overall calibration slope of 0.93 with an intercept of  $-0.29$ , indicating good calibration.

We divided the complete group into four different risk groups; low, moderate, high, and very high risk, with PRE-DELIRIC scores of 0-20%, >20-40%, >40-60%, and >60%. Figure 2⇓ shows the sensitivity, specificity, and likelihood ratios for each risk group. Figure 3⇓ shows the calibration plot of the pooled data.

### Delirium prediction by caregivers

In a convenience sample of 124 patients, we asked attending intensive care nurses and physicians to predict delirium, independently of each other, within 24 hours of patients' admission to intensive care. The AUROC for prediction by the nurses (0.59, 0.49 to 0.70) and physicians (0.59, 0.49 to 0.70) was inferior to the predictive value of the PRE-DELIRIC model (0.87, 0.81 to 0.93) in this specific subgroup of 124 patients. We found no significant differences between the prediction made by intensive care nurses (75% of sample) and student intensive care nurses (25%) or between predictions made by intensivists (36%), fellow-intensivists (40%), or residents (24%) (data not shown).

## Discussion

In this multicentre study, we developed and validated a model for predicting delirium in intensive care patients. To our knowledge, this is the first delirium prediction study for general intensive care patients and represents by far the largest delirium related study in intensive care patients to date. Our PRE-DELIRIC model reliably predicted the development of delirium for the complete length of stay in intensive care, on the basis of 10 readily available risk factors within 24 hours of admission to intensive care. In addition, the area under the receiver operating characteristics curve of the PRE-DELIRIC model was significantly higher than the delirium prediction capacity of attending caregivers. These findings confirm that the model has additional value in daily practice. Importantly, dementia and alcohol misuse are not included the model, as

**Formula for PRE-DELIRIC model**

Risk of delirium =  $1/(1+\exp(-(-6.31$   
 $+ 0.04 \times \text{age}$   
 $+ 0.06 \times \text{APACHE-II score}$   
 $+ 0 \text{ for non-coma or } 0.55 \text{ for drug induced coma or } 2.70 \text{ for miscellaneous coma or } 2.84 \text{ for combination coma}$   
 $+ 0 \text{ for surgical patients or } 0.31 \text{ for medical patients or } 1.13 \text{ for trauma patients or } 1.38 \text{ for neurology/neurosurgical patients}$   
 $+ 1.05 \text{ for infection}$   
 $+ 0.29 \text{ for metabolic acidosis}$   
 $+ 0 \text{ for no morphine use or } 0.41 \text{ for } 0.01\text{-}7.1 \text{ mg/24 h morphine use or } 0.13 \text{ for } 7.2\text{-}18.6 \text{ mg/24 h morphine use or } 0.51 \text{ for } >18.6 \text{ mg/24 h morphine use}$   
 $+ 1.39 \text{ for use of sedatives}$   
 $+ 0.03 \times \text{urea concentration (mmol/L)}$   
 $+ 0.40 \text{ for urgent admission}))$

The scoring system's intercept is expressed as  $-6.31$ ; the other numbers represent the shrunken regression coefficients (weight) of each risk factor.

these patients need to be considered as high risk patients irrespective of the presence of other risk factors.

**Clinical relevance**

The early prediction of development of delirium in intensive care patients with the PRE-DELIRIC model facilitates the use of non-drug preventive measures in high risk patients, such as improvement of orientation, cognitive stimulation, early mobilisation,<sup>11</sup> and listening to music.<sup>24</sup> It also facilitates drug interventions in high risk patients, such as the administration of prophylactic haloperidol.<sup>12</sup> These interventions aim to improve patients' cognition or have a systemic effect, although the evidence of beneficial preventive measures with drugs and nursing interventions in critically ill patients is limited at this moment.<sup>25</sup> Non-drug preventive measures were successful in reducing the incidence and duration of delirium in a non-critically ill hospital population with an intermediate to high risk for the development of delirium,<sup>11</sup> and prevention with haloperidol resulted in reduced severity of delirium and fewer days with delirium, as well as a shorter length of stay in hospital.<sup>12</sup> Importantly, no data from intensive care patients are available. Interestingly, early mobilisation of mechanically ventilated patients in intensive care, besides other significant effects, resulted in a reduced duration of delirium.<sup>26</sup>

The use of the PRE-DELIRIC model to identify and consequently preventively treat high risk patients could offer an important contribution to intensive care practice and ensure efficient use of research resources to study only high risk patients. In addition, the modifiable risk factors of the model may facilitate the use of preventive measures. Currently, the PRE-DELIRIC model is used in clinical daily practice in the hospital that developed the model; intensive care patients with a high risk of delirium ( $\geq 50\%$  PRE-DELIRIC score), and patients with dementia or alcohol misuse, receive preventive measures. The optimal cut-off point of the PRE-DELIRIC model and the most effective delirium preventive interventions for intensive care patients need to be studied in the near future.

**Limitations of study**

Our study had several limitations. Firstly, although the CAM-ICU has a high sensitivity and specificity when used by dedicated research nurses,<sup>27,28</sup> its performance in daily practice as used by bedside nurses recently proved to be lower.<sup>29</sup> In this performance study, the CAM-ICU was measured at one point on one day, whereas our diagnosis of delirium was based on all CAM-ICU screenings during patients' complete stay in intensive care, increasing the sensitivity of the test. We also used

haloperidol as a proxy for the diagnosis of delirium, as in all participating centres haloperidol was used only to treat delirium, and the hospitals with the highest CAM-ICU performance participated in this delirium prediction study. In view of the fluctuating nature of delirium, all patients were screened three times daily and more often if needed. When delirium was not detected with the CAM-ICU but suspected on the basis of medical and nursing reports, patients were additionally screened by a delirium expert according to the DSM-IV criteria.<sup>1</sup> In addition, during the development and temporal validation study, we did quality checks that showed a high compliance rate and inter-rater reliability. We therefore presume that few patients were misdiagnosed.

Secondly, we used data collected from four other hospitals in the same study period. These centres implemented and clinically used the CAM-ICU combined with a delirium treatment protocol before the conduct of the study. For the external validation study, we included only patients with complete CAM-ICU screenings and those who were treated with haloperidol for delirium. The case mix of these patients showed a higher APACHE-II score and more sedated patients, and more patients were admitted for medical reasons compared with the hospital where the primary development and validation studies were done. These differences may explain the higher incidence of delirium in these hospitals. Because of logistic reasons and the fact that we wanted to examine the predictive value of the PRE-DELIRIC model in daily intensive care practice, we did not do quality checks such as inter-rater reliability measurements in these other hospitals. Despite these limitations, the PRE-DELIRIC model showed a good predictive value in daily intensive care practice.

Thirdly, as recommended,<sup>21</sup> the risk factors used in our study were primarily based on a systematic review.<sup>18</sup> We included additional variables following the results of our first cohort. We added "diagnosis group" and "urgent admission" as new risk factors because of a high incidence of delirium associated with these items. Although these variables were not found in the systematic review,<sup>18</sup> some studies show that urgent admission to intensive care and neurological conditions are risk factors for delirium.<sup>30,31</sup> The results of our development study show that these risk factors are of importance in predicting delirium in intensive care patients. Because of a low prevalence rate, relevant risk factors such as hyponatraemia, alcohol misuse, and dementia were excluded from the multivariate logistic regression analysis. The additional value of hyponatraemia for the model would be expected to be low, as the incidence of delirium in patients with hyponatraemia in the first 24 hours after admission to intensive care is low. The importance of, for

example, dementia and alcohol misuse is recognised in several studies,<sup>4 32</sup> and the incidence of delirium in these patients was also high in our study. In many institutes, all these patients will receive preventive measures so physicians do not need a delirium prediction model in these particular subgroups. Moreover, adding these covariates to the model would decrease its sensitivity to the other covariates. For these reasons, we did not include alcohol misuse and dementia in the PRE-DELIRIC model.

Fourthly, the negative likelihood ratio for patients with a predicted low chance of developing delirium is relatively rather moderate. This indicates that, in this group, patients will develop delirium while they are classified as having a low risk. On the other hand, preventive measures are advised in patients with a high risk, and the higher the risk of delirium the better the performance of the model. Nevertheless, a predicted low risk does not exclude the possibility of development of delirium.

Finally, the PRE-DELIRIC model is a static model that yields a calculated probability for delirium 24 hours after admission to intensive care. As the health status of patients can improve or deteriorate during their stay in intensive care, the probability of development of delirium may also change. Our model does not take into account changes in health status. Despite this limitation of the PRE-DELIRIC model, the area under the receiver operating characteristics curve of the model is high. Even so, development of a dynamic prediction model using dynamic parameters, such as the sequential organ failure assessment (SOFA) score would be interesting, to improve its predictive value during the patients' stay on the intensive care unit, which may also result in a better performance in the low risk group.

## Conclusions and policy implications

The PRE-DELIRIC model can predict delirium for the complete stay in the intensive care within 24 hours of admission. We can now identify patients who have a high risk of developing delirium during their intensive care stay. This will facilitate targeted initiation of preventive measures. Our study shows that the use of the PRE-DELIRIC model is significantly better than the predictions of the attending caregivers, and it should therefore be used daily in intensive care practice.

An automatic version of the PRE-DELIRIC model (Excel and web based) can be downloaded at [www.umcn.nl/Research/Departments/intensive%20care/Pages/vandenBoogaard.aspx](http://www.umcn.nl/Research/Departments/intensive%20care/Pages/vandenBoogaard.aspx) (English and Dutch version available).

We thank Gabriel Roodbol, delirium expert psychiatric nurse, who did all the inter-rater reliability measurements during the study. Thanks also to Henk Westerbaan, Maaike Fenten, and Wendy Groetelaers-Kusters, for collecting data during the absence of the first author and processing the data in SPSS. We also thank Sjeff van der Velde for his support in gathering the national intensive care evaluation database information. Finally, we thank Gareth Parsons (University of Glamorgan, Faculty of Health, Sport and Science, UK) for his support with English grammar and spelling.

Contributors: MvdB conducted the study, collected all data, did the statistical analysis in collaboration with RD, and drafted the manuscript. PP and LS supervised the conduct of the study and writing of the paper. AJCS, MAK, PES, and PHJvdV gathered the data from the other hospitals and corrected the manuscript. JGvdH and TvA co-supervised and corrected the manuscript. MvdB and PP are the guarantors.

Funding: None.

Competing interests: All authors have completed the Unified Competing Interest form at [http://www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on

request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: The regional medical ethical committee approved the study and waived the need for informed consent, as no additional interventions were carried out and data collection was not burdensome to patients.

Data sharing: No additional data available.

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-IV). 4th ed. APA, 1994.
- Dubois MJ, Bergeron N, Dumont M, Dial S, Skrobik Y. Delirium in an intensive care unit: a study of risk factors. *Intensive Care Med* 2001;27:1297-304.
- Ely EW, Gautam S, Margolin R, Francis J, May L, Speroff T, et al. The impact of delirium in the intensive care unit on hospital length of stay. *Intensive Care Med* 2001;27:1892-900.
- Quimet S, Kavanagh BP, Gottfried SB, Skrobik Y. Incidence, risk factors and consequences of ICU delirium. *Intensive Care Med* 2007;33:66-73.
- Milbrandt EB, Deppen S, Harrison PL, Shintani AK, Speroff T, Stiles RA, et al. Costs associated with delirium in mechanically ventilated patients. *Crit Care Med* 2004;32:955-62.
- Luetz A, Heymann A, Radtke FM, Chenitir C, Neuhaus U, Nachtigall I, et al. Different assessment tools for intensive care unit delirium: which score to use? *Crit Care Med* 2010;38:409-18.
- Van Eijk MJ, van Marum RJ, Klijn I, de Wit N, Kesecioglu J, Slooter AJC. Comparison of delirium assessment tools in a mixed intensive care unit. *Crit Care Med* 2009;37:1881-5.
- Koster S, Hensens AG, Oosterveld FG, Wijma A, van der Palen J. The delirium observation screening scale recognizes delirium early after cardiac surgery. *Eur J Cardiovasc Nurs* 2009;8:309-14.
- Spronk PE, Riekerk B, Hofhuis J, Rommes JH. Occurrence of delirium is severely underestimated in the ICU during daily care. *Intensive Care Med* 2009;35:1276-80.
- Van den Boogaard M, Pickkers P, van der Hoeven H, Roodbol G, van Achterberg T, Schoonhoven L. Implementation of a delirium assessment tool in the ICU can influence haloperidol use. *Crit Care* 2009;13:R131.
- Inouye SK, Bogardus ST Jr, Charpentier PA, Leo-Summers L, Acampora D, Holford TR, et al. A multicomponent intervention to prevent delirium in hospitalized older patients. *N Engl J Med* 1999;340:669-76.
- Kalisvaart KJ, de Jonghe JF, Bogaards MJ, Vreeswijk R, Egberts TC, Burger BJ, et al. Haloperidol prophylaxis for elderly hip-surgery patients at risk for delirium: a randomized placebo-controlled study. *J Am Geriatr Soc* 2005;53:1658-66.
- Inouye SK, Viscoli CM, Horwitz RJ, Hurst LD, Tinetti ME. A predictive model for delirium in hospitalized elderly medical patients based on admission characteristics. *Ann Intern Med* 1993;119:474-81.
- Inouye SK, Zhang Y, Jones RN, Kiely DK, Yang F, Marcantonio ER. Risk factors for delirium at discharge: development and validation of a predictive model. *Arch Intern Med* 2007;167:1406-13.
- Pisani MA, Murphy TE, Van Ness PH, Araujo KL, Inouye SK. Characteristics associated with delirium in older patients in a medical intensive care unit. *Arch Intern Med* 2007;167:1629-34.
- Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000;19:453-73.
- Vreeswijk R, Toornvliet A, Honing M, Bakker K, de Man T, Daas G, et al. Validation of the Dutch version of the confusion assessment method (CAM-ICU) for delirium screening in the intensive care unit. *Netherlands Journal of Critical Care* 2009;13:73-8.
- Van Rompaey B, Schuurmans MJ, Shortridge-Baggett LM, Truijten S, Bossaert L. Risk factors for intensive care delirium: a systematic review. *Intensive Crit Care Nurs* 2008;24:98-107.
- Arts D, de Keizer NF, Scheffer GJ, de Jonge E. Quality of data collected for severity of illness scores in the Dutch National Intensive Care Evaluation (NICE) registry. *Intensive Care Med* 2002;28:656-9.
- Inouye SK, Leo-Summers L, Zhang Y, Bogardus ST Jr, Leslie DL, Agostini JV. A chart-based method for identification of delirium: validation compared with interviewer ratings using the confusion assessment method. *J Am Geriatr Soc* 2005;53:312-8.
- Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. Springer Science+Business Media, 2009.
- R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, 2009.
- Harrell FE Jr, with contributions from many other users. The Hmisc package. Harrell Miscellaneous, 2009.
- McCaffrey R, Locsin R. The effect of music on pain and acute confusion in older adults undergoing hip and knee surgery. *Holist Nurs Pract* 2006;20:218-24.
- Devlin JW, Skrobik Y. Antipsychotics for the prevention and treatment of delirium in the intensive care unit: what is their role? *Harv Rev Psychiatry* 2011;19:59-67.
- Schweickert WD, Pohlman MC, Pohlman AS, Nigos C, Pawlik AJ, Esbrook CL, et al. Early physical and occupational therapy in mechanically ventilated, critically ill patients: a randomised controlled trial. *Lancet* 2009;373:1874-82.
- Ely EW, Inouye SK, Bernard GR, Gordon S, Francis J, May L, et al. Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU). *JAMA* 2001;286:2703-10.
- Ely EW, Margolin R, Francis J, May L, Truman B, Dittus R, et al. Evaluation of delirium in critically ill patients: validation of the confusion assessment method for the intensive care unit (CAM-ICU). *Crit Care Med* 2001;29:1370-9.
- Van Eijk MJ, van den Boogaard M, van Marum RJ, Benner P, Eikelenboom P, Honing ML, et al. Routine use of the confusion assessment method for the intensive care unit: a multicenter study. *Am J Respir Crit Care Med* 2011;184:340-4.
- Aldemir M, Ozen S, Kara IH, Sir A, Bac B. Predisposing factors for delirium in the surgical intensive care unit. *Crit Care* 2001;5:265-70.
- Oldenbeuving AW, de Kort PL, Jansen BP, Roks G, Kappelle LJ. Delirium in acute stroke: a review. *Int J Stroke* 2007;2:270-5.

**What is already known on this topic**

Delirium is a common and serious disorder related to morbidity and mortality

Several risk factors for delirium in intensive care patients are recognised, but no delirium prediction model is available

**What this study adds**

A delirium prediction model has been developed and validated

The PRE-DELIRIC model enables the physician to predict delirium within 24 hours after admission to intensive care for the complete stay in intensive care

This model may facilitate early identification of high risk patients and targeted early initiation of preventive measures

32 McNicoll L, Pisani MA, Zhang Y, Ely EW, Siegel MD, Inouye SK. Delirium in the intensive care unit: occurrence and clinical course in older patients. *J Am Geriatr Soc* 2003;51:591-8.

**Accepted:** 14 November 2011

Cite this as: *BMJ* 2012;344:e420

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license. See: <http://creativecommons.org/licenses/by-nc/2.0/> and <http://creativecommons.org/licenses/by-nc/2.0/legalcode>.

## Tables

Table 1 | Characteristics of participating hospitals

| Participating hospitals                               | ICU beds for adults (annual admissions) | ICU population  | CAM-ICU implementation              | CAM-ICU screening                            |
|---|---|---|-------------------------------------|--|
| Radboud University Nijmegen Medical Centre, Nijmegen* | 33 beds (2000-2500)                     | Medicine, surgery, neurocritical care, and cardiothoracic surgery | 2007; group and individual training | 3/day; compliance rate 90%; IRR>0.8          |
| University Medical Centre Utrecht, Utrecht*           | 32 beds (2000-2500)                     | Medicine, surgery, neurocritical care, and cardiothoracic surgery | 2007; group and individual training | 2/day; compliance and IRR not measured       |
| Medical Centre Leeuwarden, Leeuwarden†                | 16 beds (1400-1500)                     | Medicine, surgery, and cardiothoracic surgery                     | 2008; group and individual training | 3/day; compliance rate 90%; IRR not measured |
| Gelre Hospital, Apeldoorn†                            | 10 beds (600)                           | Medicine and surgery  | 2004; group and individual training | 3/day; compliance rate 90%; IRR not measured |
| Onze Lieve Vrouwe Gasthuis, Amsterdam†                | 18 beds (1500-1800)                     | Medicine, surgery, and cardiothoracic surgery                     | 2006; group and individual training | 3/day; compliance rate 96%; IRR not measured |

CAM=confusion assessment method; ICU=intensive care unit; IRR=inter-rater reliability expressed as Cohen's  $\kappa$ .

\*University hospital.

†University affiliated teaching hospital.

Table 2| Patients' characteristics in cohort studies. Values are numbers (percentages) unless stated otherwise

| Variable  | Development study<br>(n=1613) | Temporal validation study<br>(n=549) | External validation study<br>(n=894) | All included patients (n=3056) |
|---|-------------------------------|--------------------------------------|--------------------------------------|--------------------------------|
| Median (IQR, range) age (years)                       | 64 (54-72, 76)                | 64 (54-73, 72)                       | 67 (58-75, 78)                       | 65 (55-73, 78)                 |
| Male sex  | 1027 (63.7)                   | 353 (64.3)                           | 557 (62.3)                           | 1937 (63.4)                    |
| Median (IQR, range) APACHE-II score                   | 14 (11-18, 41)                | 15 (11-18, 47)                       | 16 (13-21, 44)                       | 15 (12-19, 48)                 |
| Delirium  | 411 (25.5)                    | 171 (31.1)                           | 329 (36.8)                           | 911 (29.8)                     |
| Median (IQR, range) days to onset of delirium         | 2 (2-5, 53)                   | 2 (2-5, 44)                          | NA                                   | NA                             |
| Urgent admission                                      | 852 (52.8)                    | 232 (42.3)                           | 495 (55.4)                           | 1579 (51.7)                    |
| Mechanical ventilation                                | 1266 (78.5)                   | 478 (87.1)                           | 801 (89.6)                           | 2545 (83.3)                    |
| Sedation  | 386 (23.9)                    | 184 (33.5)                           | 543 (60.7)                           | 1113 (36.4)                    |
| Median (IQR, range) length of stay in ICU (days)      | 2 (2-4, 118)                  | 2 (2-4, 69)                          | 4 (2-8, 100)                         | 3 (2-7, 118)                   |
| Median (IQR, range) length of stay in hospital (days) | 9 (5-19, 247)                 | 9 (5-15, 98)                         | 10 (6-20, 88)                        | 10 (5-19, 249)                 |
| Admission category:                                   |                               |                                      |                                      |                                |
| Surgery   | 1010 (62.6)                   | 340 (61.9)                           | 507 (56.7)                           | 1857 (60.8)                    |
| Medical   | 360 (22.3)                    | 123 (22.4)                           | 297 (33.2)                           | 780 (25.5)                     |
| Trauma  | 80 (5.0)                      | 18 (3.3)                             | 30 (3.4)                             | 128 (4.2)                      |
| Neurology/neurosurgery                                | 163 (10.1)                    | 68 (12.4)                            | 60 (6.7)                             | 291 (9.5)                      |

ICU=intensive care unit; IQR=interquartile range; NA=not applicable.

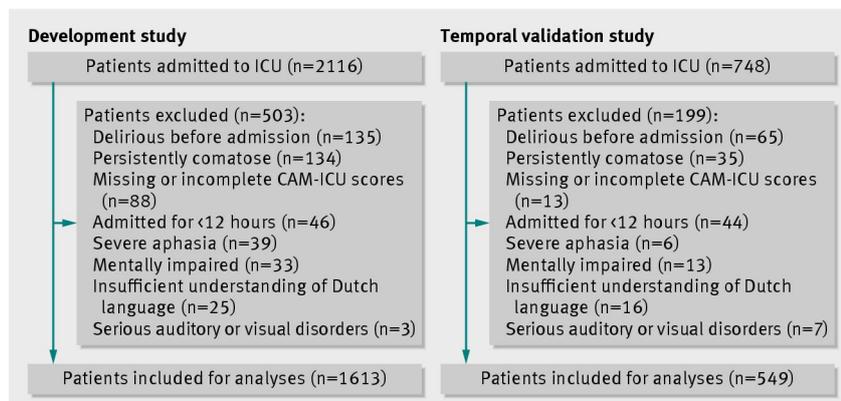
Table 3| Variables of PRE-DELIRIC model and regression coefficients

| Variable                    | Regression coefficient | Odds ratio (95% CI) | Shrunken regression coefficient* |
|-----------------------------|------------------------|---------------------|----------------------------------|
| Age (years)                 | 0.04                   | 1.04 (1.03 to 1.06) | 0.04                             |
| APACHE-II score (per point) | 0.06                   | 1.06 (1.03 to 2.0)  | 0.06                             |
| Coma:                       |                        |                     |                                  |
| Drug induced                | 0.59                   | 1.8 (1.1 to 3.1)    | 0.55                             |
| Miscellaneous               | 2.92                   | 18.5 (4.6 to 73.8)  | 2.70                             |
| Combination                 | 3.06                   | 21.3 (5.9 to 77.1)  | 2.84                             |
| Admission category:         |                        |                     |                                  |
| Surgery                     | Reference              | Reference           | Reference                        |
| Medical                     | 0.33                   | 1.4 (0.9 to 2.2)    | 0.31                             |
| Trauma                      | 1.22                   | 3.4 (1.7 to 6.8)    | 1.13                             |
| Neurology/neurosurgery      | 1.49                   | 4.5 (2.6 to 7.5)    | 1.38                             |
| Infection (yes)             | 1.14                   | 3.1 (2.0 to 4.8)    | 1.05                             |
| Metabolic acidosis (yes)    | 0.32                   | 1.4 (1.0 to 2.0)    | 0.29                             |
| Morphine use:               |                        |                     |                                  |
| 0.01-7.1 mg/day             | 0.44                   | 1.6 (0.8 to 3.1)    | 0.41                             |
| 7.2-18.6 mg/day             | 0.14                   | 1.2 (0.8 to 1.8)    | 0.13                             |
| >18.6 mg/day                | 0.55                   | 1.8 (1.1 to 2.7)    | 0.51                             |
| Sedation (yes)              | 1.51                   | 4.5 (2.8 to 7.4)    | 1.39                             |
| Urea (mmol/L)               | 0.03                   | 1.03 (1.0 to 1.1)   | 0.03                             |
| Urgent admission (%)        | 0.43                   | 1.5 (1.1 to 2.3)    | 0.40                             |
| Intercept                   | -6.76                  |                     | -6.31                            |

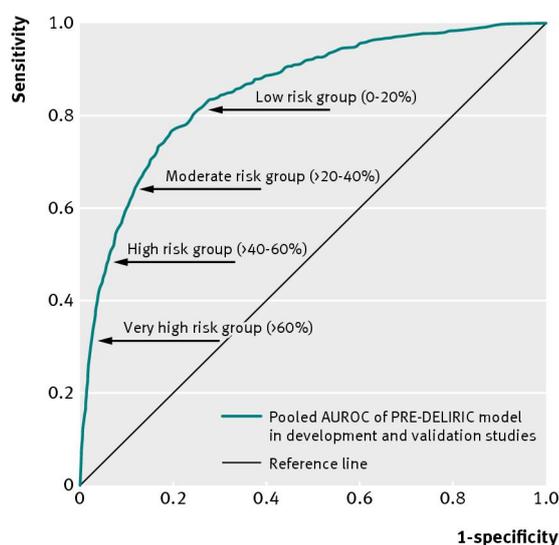
APACHE=acute physiology and chronic health evaluation.

\*Result of bootstrapping technique used to correct for overoptimistic estimation of model.

## Figures

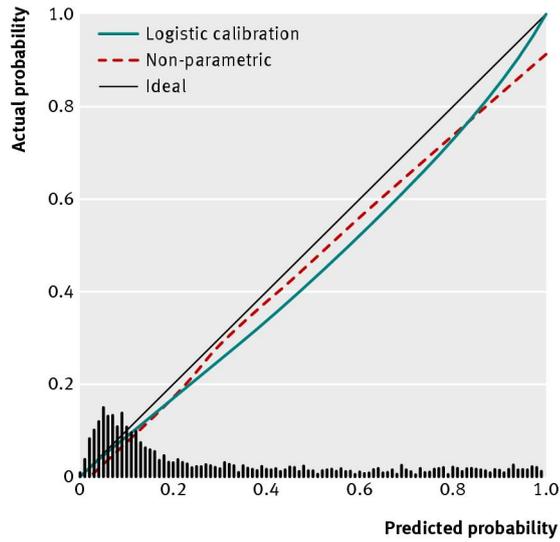


**Fig 1** Flow chart of development and temporal validation study



|                | Sensitivity (%) | Specificity (%) | Positive likelihood ratio | Negative likelihood ratio |
|----------------|-----------------|-----------------|---------------------------|---------------------------|
| Low risk       | 80.7            | 74.7            | 3.19                      | 0.26                      |
| Moderate risk  | 62.0            | 88.8            | 5.54                      | 0.43                      |
| High risk      | 46.3            | 94.6            | 8.57                      | 0.57                      |
| Very high risk | 30.0            | 97.7            | 13.0                      | 0.72                      |

**Fig 2** Area under receiver operating characteristics curve (AUROC) of pooled data for development (AUROC=0.86), temporal validation (AUROC=0.89), and external validation (AUROC=0.84), resulting in AUROC of 0.85. AUROC was 0.59 (95% CI 0.49 to 0.70) for prediction by nurses and 0.59 (0.49 to 0.70) for prediction by physicians



**Fig 3** Calibration plot of pooled data, with calibration slope of 0.93 and intercept of  $-0.29$