BMJ Open

# Can mid-regional pro-adrenomedullin (MR-proADM) increase the prognostic accuracy of NEWS in predicting deterioration in patients admitted to hospital with mild to moderately severe illness? A prospective single-centre observational study.



| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2017-020337 |
| Article Type: | Research |
| Date Submitted by the Author: | 30-Oct-2017 |
| Complete List of Authors: | Graziadio, Sara; Newcastle Upon Tyne Hospitals NHS Foundation Trust, NIHR Diagnostic Evidence Co-operative<br>Price, David; Newcastle Upon Tyne Hospitals NHS Foundation Trust, NIHR Diagnostic Evidence Co-operative<br>O'Leary, Rachel; Newcastle University, NIHR Diagnostic Evidence Co-operative<br>Stocken, Deborah; Newcastle University, Newcastle Clinical Trial Unit, Institute of Health and Society,<br>Power, Michael; Newcastle Upon Tyne Hospitals NHS Foundation Trust, NIHR Diagnostic Evidence Co-operative; Newcastle University, NIHR Diagnostic Evidence Co-operative<br>Allen, A; Newcastle University, NIHR Diagnostic Evidence Co-operative<br>Simpson, A John; Newcastle University Institute of Cellular Medicine, Institute of Cellular Medicine, Medical School |
| Primary Subject Heading: | Diagnostics |
| Secondary Subject Heading: | Evidence based practice |
| Keywords: | INTERNAL MEDICINE, ACCIDENT & EMERGENCY MEDICINE, Adult intensive & critical care < INTENSIVE & CRITICAL CARE |
| | |

Prognostic accuracy of MR-proADM in emergency departments

**Title:** Can mid-regional pro-adrenomedullin (MR-proADM) increase the prognostic accuracy of NEWS in predicting deterioration in patients admitted to hospital with mild to moderately severe illness? A prospective single-centre observational study.

**Authors:** Sara Graziadio[1], D. Ashley Price[2], Rachel A. O'Leary [1], Deborah D. Stocken[3], Michael Power[1], A. Joy Allen[1], A. John Simpson[1]

[1] National Institute for Health Research (NIHR) Diagnostic Evidence Cooperative (DEC) Newcastle, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK

[2] Department of Infectious Diseases, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, NE1 4LP, UK

[3] Newcastle Clinical Trials Unit, Newcastle University, 1-4 Claremont Terrace, Newcastle upon Tyne, NE2 4AE, UK

**Corresponding author:** Dr Sara Graziadio, NIHR Diagnostic Evidence Co-operative Newcastle, M2.081 William Leech Building, Medical School, Newcastle University, Newcastle upon Tyne, NE2 4HH, U.K. sara.graziadio@ncl.ac.uk

Word counts

Abstract: 282

Paper: 3104

Prognostic accuracy of MR-proADM in emergency departments


## Abstract

**Objective** To assess the accuracy of a statistical model based on admission NEWS score and MR-proADM blood level in predicting deterioration in mild to moderately ill people.

**Design** Prospective observational study

**Setting** The Medical Admissions Suite of the Royal Victoria Infirmary, Newcastle.

**Participants** 300 adults with NEWS score between 2 and 5 on admission. Exclusion criteria included receiving palliative care, or admitted for social reasons or self-harming. Patients were enrolled between September and December 2015, and followed-up for 30 days after discharge.

**Outcome measure** The primary outcome measure was the proportion of patients who, within 72 hours, had an *Acuity Increase*, defined as any combination of: an increase of at least 3 in the NEWS score; transferred to a higher-dependency bed or monitored area; and, for those discharged from hospital, re-admission for medical reasons; or death.

**Results** NEWS predicted *Acuity Increase* poorly: the area under the curve (AUC) was 0.55 (95% CI 0.48, 0.62) with univariate analysis. NEWS and MR-proADM together predicted *Acuity Increase* more accurately, increasing AUC to 0.61 (95% CI 0.54, 0.69). When the presence of chronic obstructive pulmonary disease or heart failure and interaction with MR-proADM were added to the model, the predictive accuracy further increased the AUC to 0.69 (95% CI 0.63, 0.76).

**Conclusions** MR-proADM improves the accuracy of prediction by NEWS of deterioration in patients admitted to hospital with a mild to moderately severe acute illness. As a growing number of NHS hospitals are implementing the NEWS score on their clinical information systems, further research should assess the practicalities and utility of developing a decision aid based on admission NEWS score, MR-proADM level, and possibly other clinical data and other biomarkers that could further improve prediction accuracy.



## Keywords

Biochemistry, diagnosis, health services research

Prognostic accuracy of MR-proADM in emergency departments

## Strengths and limitations of this study

- This is the first study to use rigorous statistical methods to assess the value added by MR-proADM to the admission NEWS score for predicting clinically important deterioration in mild to moderately ill patients
- Prediction accuracy might have been greater had more severely ill patients been included, but these people are already known to be severely ill.
- This was an observational study, and thus could not directly assess the utility of more accurate prediction of deterioration

## Introduction

The National Early Warning Score (NEWS) is recommended for assessing severity of illness in patients presenting in primary or secondary NHS care and for surveillance of patients in hospital [1 2]. Six physiological parameters (which can be measured at the bedside) are scored: respiratory rate, oxygen saturation, temperature, systolic blood pressure, pulse rate, and level of consciousness. The scores are aggregated, and, if the patient requires oxygen, the total is increased. NEWS predicts death, cardiac arrest, and unplanned intensive care unit (ICU) admission within 24 hours [3-5]. However, NEWS does not identify all patients who turn out to be seriously ill [6-8], and there are also patients whose NEWS score is usually elevated and who do not require the level of observation that the NEWS tool would suggest. For example, people with chronic obstructive pulmonary disease (COPD) or chronic heart failure (HF) have higher baseline NEWS scores than those without these comorbidities. The useful predictive accuracy of NEWS for patients presenting to the Emergency Department (ED) has been confirmed in a wide range of severity of illness [9 10], as has its reduced accuracy in people with COPD [11]. But, no previous studies of the predictive accuracy of NEWS in the ED/Medical Admissions Unit (MAU) have focussed on patients admitted with mild to moderately severe illness. Since a clinically important proportion of these patients do deteriorate unexpectedly, improved risk stratification would be useful.

Mid-regional pro-adrenomedullin (MR-proADM) is one of several promising biomarkers for severe illness and deterioration [12-16].

MR-proADM is a precursor of adrenomedullin (ADM), a member of the calcitonin peptide family. ADM is widely expressed and has roles in vasodilation, immune modulation, and metabolic regulation. It is up-regulated in severe infections, inflammation, vasodilation, stimulation of diuresis, increased cardiac output, and stroke [17-19]. ADM has a short half-life, but MR-proADM is more stable and directly reflects ADM concentrations in blood. Both ADM and MR-proADM levels are strongly associated with risk of mortality, regardless of aetiology [20-26] . In people presenting with acute chest pain, MR-proADM has been reported to improve the Global Registry of Acute Coronary Events risk classification by 41% [27]. As with the NEWS score, people with COPD or chronic heart failure have higher baseline levels of MR-proADM.

Prognostic accuracy of MR-proADM in emergency departments

The aim of this study was to assess whether the MR-proADM level used alongside the NEWS score would improve prediction of deterioration over NEWS score alone in patients admitted to the MAU with mild to moderately severe illness.

# Methods

## Study participants and study design

This was a prospective observational cohort study. Patients were enrolled between September and December 2015 at the Royal Victoria Infirmary, Newcastle, and followed-up for 30 days after discharge. If the patient died within the 30 days of follow up, this and cause of death were recorded. Adults admitted to the MAU were recruited for the study between 9am and 4pm on weekdays.

Sample size was determined as a pragmatic recruitment target for a three-month observational study. A recent unpublished audit conducted in the MAU at the Royal Victoria Infirmary found a deterioration rate of 20%. With 300 patients and complete data collection, 60 events would be anticipated. With this number of events, a multivariable prediction model could reliably include up to six independent predictors: models with fewer than ten events per predictor tend to be over-fitted [28].

Patients were considered eligible for inclusion in the study if their NEWS score on admission was at least 2 and not greater than 5, and all NEWS parameters were recorded. Patients were excluded from the study if they were receiving palliative care, were admitted for social reasons only, or were self-harming, or overdosing with drugs or other substances.

All participants provided written informed consent, and the study was approved by the Newcastle & North Tyneside Research Ethics Committee (15/NE/0120).

## Recorded data

Demographic and admission data included: gender, year of birth, reason for admissions, diagnosis on discharge, and the presence of comorbidities in which baseline MR-proADM levels are chronically raised: COPD with hypoxia ($PaO_2$ <10 kPa) [7]; HF [29]; acute brain injury [6]; acute coronary syndrome [27]; acute venous thromboembolism [21]; high International Normalized Ratio (INR>2); acute kidney injury; electrolyte disturbances

(Na⁺ <130 or >150 mmol/L; K⁺ <3.0 or >5.5 mmol/L); hyperglycaemia in type 1 diabetes (random glucose >10 mmol/L).

The NEWS score was assessed at baseline and over the next 72 hours, and the scores and assessment times recorded. The 7 clinical parameters used to determine the NEWS score were recorded for the baseline assessment only.

Blood samples were taken at hospital admission for assessment of MR-proADM, C-Reactive Protein (CRP) and white blood count (WBC).

## Laboratory tests

Plasma was obtained from blood samples (collected in ethylenediaminetetraacetic acid, EDTA) that were no longer clinically required. Plasma was stored in aliquots at –80º C

MR-proADM was assayed in the on-site Blood Sciences Laboratory using the BRAHMS Kryptor system according to the manufacturer's instructions.

Blood samples were analysed in batches by personnel blinded with regard to the condition and NEWS score of the patient. Nurses who assessed the NEWS score and healthcare professionals managing patients in the MAU were blinded to MR-proADM results.

## Outcomes of interest

**Outcome 1: *Acuity Increase* (i.e. deterioration)**. A patient was classified as having an *Acuity Increase* if one or more of the following occurred within 72 hours from admission:

1. transfer to a higher level of care (ICU or high dependency unit)
2. readmission to hospital for reasons related to the initial admission
3. death for reasons related to the initial admission
4. NEWS score increased by at least two compared to the admission score

**Outcome 2: *Deterioration Event***. For most of the observed *Acuity Increase* cases the reason for classification was an increase in the NEWS score (Table 1). Because an increase in NEWS score reflects both measurement variation and physiological variation, additional exploratory analyses were carried out to assess the performance of MR-proADM in predicting deterioration. *Deterioration Events* were classified as the occurrence of one or more of the following:

1. transfer to higher level of care within 72 hours from admission;
2. death for related reason to admission within 30 days;
3. re-admission to hospital (for the same reason as the previous admission) within 30 days from first admission.

Classification based on this definition is unlikely to be subject to clinically important measurement variation. This analysis, therefore, should optimise the accuracy of prediction accuracy for events which are both clinically and economically important.

**Outcome 3: *Length of Stay***. *Length of Stay* was defined as the duration (in days) from admission to discharge or death.

## Statistical analysis

All data analyses were performed using the R language version 3.2.0 [30], with the support of RStudio, version 0.99.896 (RStudio, Inc). The following R packages were used: ggplot2, pROC, psych, PredictABEL, Hmisc. For the ROC curve analyses, data were exported to SPSS version 22 (SPSS, Inc., Chicago IL) and analyses were re-run for quality assurance of results.

Logistic regression models were compared for their accuracy in predicting deterioration outcome measures as pre-specified in the analysis plan. Analyses are presented as unadjusted parameter estimates of risk (odds ratio (OR), with confidence intervals) and estimates adjusted for identified clinical confounding factors. The aims of the multivariable analyses were twofold: first, estimate the effect size and significance adjusted for other identified influential predictors and interactions; second, to investigate whether the addition of other predictors improved the goodness of fit and accuracy of prediction.

Only complete cases were analysed since missingness was minimal (see Table 1).

For each outcome of deterioration, logistic regression models were compared for the following sets of predictor variables:

*Predictor set a.* Comparator (base case): NEWS score on admission

*Predictor set b.* Primary analysis: NEWS score, MR-proADM

*Predictor set c.* Secondary analyses: NEWS score and MR-proADM always included. Age, gender, CRP, WBC, presence of COPD or HF,

presence of other comorbidities, and interactions between predictors when appropriate.

Predictors (and the underlying assumption of linearity of their relationship with the outcome of interest) were initially investigated through univariate analyses based on simple log and quadratic functions. We have assessed interactions through visual data exploration. Transformations were applied if they improved the goodness of fit as assessed by the Akaike information criterion (AIC), and were retained in the multivariable setting. Subsequently, for the multivariate regression the set of predictors was reduced through backward elimination, again based on changes in AIC.

Secondary outcome of *Length of Stay* followed a similar analysis plan using multiple regressions based on a transformed outcome to address non-normality. To evaluate whether MR-proADM was a predictor of the length of stay in the hospital, linear regression was used since the outcome variable (*Length of Stay*) is a continuous variable. Variables were log-transformed if not normally distributed. Normality was assessed by visualizing the data. The regression model included *Length of Stay* as outcome variable, and NEWS and MR-proADM as predictors. More details of the methods used are reported in the Supplemental Data.

Predictive accuracy of the models was assessed with the area under the ROC curve (AUC) and is presented for all models with 95% confidence intervals (CI). To assess the value added by including the MR-proADM level with the NEWS score in predicting deterioration, continuous net reclassification improvement (NRI) and integrated discrimination improvement (IDI) were calculated[31 32].

Internal validation of models was performed through bootstrapping with 10,000 resamples.

# Results

## Study enrolment

The process of recruitment and enrolment of patients for the study is shown in Figure 1. The study recruited 300 patients, and the 292 were included in the analysis. Five patients were excluded because the blood samples for MR-proADM were taken more

than 12 hours from baseline NEWS assessment; 3 patients were excluded from the primary outcome due to missing follow up NEWS scores.

## Patient characteristics

Patient demographics and mean biomarker levels for each covariate are reported in Table 1. The cohort was evenly divided in gender and had a mean age of 63 years and mean NEWS on admission of 3, with the majority of patients having NEWS score of 2. COPD or HF were present in 28%, and 25% had other comorbidities.

Prognostic accuracy of MR-proADM in emergency departments

**Table 1**. Characteristics of the study population, classified by Outcome 1 (*Acuity Increase*), Outcome 2 (*Deterioration Event*) and All patients. Data are presented as number (no) and percentages (%) for counts, or mean and (standard deviation, SD) for continuous normally distributed data, or [25th; 50th; 75th percentile] for continuous non-normally distributed data.

| | **Outcome 1: *Acuity Increase*** | | **Outcome 2: *Deterioration Event*** | | **All patients (n = 292)** |
|---|---|---|---|---|---|
| | **Present (e = 84)** | **Absent** | **Present (e2 = 32)** | **Absent** | |
| Age (mean years, SD) | 65 (17) | 62 (21) | 63 (14) | 63 (20) | 63 (20) |
| Gender (no. females, %) | 41 (49%) | 107 (51%) | 15 (47%) | 133 (51%) | 148 (51%) |
| NEWS = 2 (no., %) | 34 (40%) | 82 (40%) | 12 (38%) | 104 (40%) | 116 (40%) |
| NEWS = 3 (no., %) | 26 (31%) | 59 (28%) | 9 (28%) | 76 (29%) | 85 (29%) |
| NEWS = 4 (no., %) | 11 (13%) | 43 (21%) | 4 (13%) | 50 (19%) | 54 (18%) |
| NEWS = 5 (no., %) | 13 (15%) | 24 (12%) | 7 (22%) | 30 (12%) | 37 (13%) |
| MR-proADM (mean nmol/l, SD) | 1.50 (1.4)<br>[0.72, 1.12, 1.79] | 1.19 (0.9)<br>[0.68, 0.93, 1.28] | 1.89 (2.0)<br>[0.93, 1.13, 1.95] | 1.20 (0.9)<br>[0.68, 0.93, 1.39] | 1.28 (1.1)<br>[0.68, 0.97, 1.48] |
| CRP (mg/l) | 59 (79)<br>[5, 22, 80] | 42 (70)<br>[4, 13, 41] | 61 (90)<br>[7, 23, 67] | 45 (71)<br>[4, 16, 51] | 47 (73)<br>[4, 17, 54] |
| WBC (x$10^9$/l) | 12 (5)<br>[9, 10, 14] | 11 (5)<br>[8, 10, 14] | 12 (4)<br>[9, 12, 15] | 11 (5)<br>[8, 10, 14] | 11 (5)<br>[8, 10, 14] |
| COPD/HF (no, %)* | 33 (39%) | 46 (22%) | 12 (38%) | 67 (26%) | 79 (28%) |
| Other comorbidities (no., %) | 17 (20%) | 55 (26%) | 15 (47%) | 57 (22%) | 72 (25%) |
| Length of stay (hrs) | 168 (196)<br>[63, 110, 194] | 137 (176)<br>[26, 68, 176] | 173 (172)<br>[59, 106, 259] | 143 (172)<br>[33, 72, 176] | 146 (182)<br>[35, 77, 182] |
| Length of stay in MAU (hrs) | 31 (19)<br>[17, 25, 43] | 24 (16)<br>[13, 21, 30] | 27 (17)<br>[18, 23, 35] | 26 (17)<br>[15, 22, 31] | 26 (17)<br>[15, 22, 31] |
| Monitored beds (no, %) | 31 (37%) | 58 (27%) | 11 (34%) | 78 (30%) | 89 (30%) |

Prognostic accuracy of MR-proADM in emergency departments

| | Outcome 1: *Acuity Increase* | | Outcome 2: *Deterioration Event* | | All patients (n = 292) |
|---|---|---|---|---|---|
| | Present (e = 84) | Absent | Present (e2 = 32) | Absent | |
| Deterioration time (hrs) | 15 (13) [5, 9, 21] | N/A | 170 (226) [19, 33, 301] | N/A | |
| * for COPD: e = number with *Acuity Increase* = 82; e2 = number with *Deterioration Event = 29;* n = total number of patients = 282 | | | | | |

211

**Table 2**. Criteria met by patients classified with an *Acuity Event* or *Deterioration Event.*

| Criterion for deterioration | *Acuity Increase* (e = 84) | *Deterioration Event* (e2 = 32) |
|---|---|---|
| NEWS (no, %) | 81 (96.4%) | N/A |
| ICU transfer (no, %) | 1 (1.2%) | 4 (12.5%) |
| Death (no, %) | 0 (0%) | 6 (18.8%) |
| Readmission (no, %) | 2 (2.4%) | 22 (68.7%) |

The study population was homogenous across *Acuity Increase* and No *Acuity Increase* outcomes in terms of gender, age, and NEWS on admission. Patients who experienced *Acuity Increase* had higher MR-proADM and CRP levels at admission, and longer length of stay in the hospital and in the MAU.

The prevalence of *Acuity Increase* was 29% (somewhat higher than the anticipated 20%). The prevalence of *Deterioration Event*s was 11%. The numbers of events provided sufficient statistical power to assess statistical significance for the primary outcome, *Acuity Increase*, but not for the secondary outcome, *Deterioration Event*, and those results should be regarded as exploratory.

## Accuracy of MR-proADM for predicting *Acuity Increase*

In the univariate analyses (Table 3) of predictors of *Acuity Increase*, the variables were transformed in a preliminary analysis assessing for non-linear relationships with the outcome variable. The final analysis used untransformed variables for all predictors except *Age*, for which a quadratic transformation, $Age^2$, was used. Potentially useful predictors of *Acuity Increase* were MR-proADM (OR = 1.27, 95% CI 1.02, 1.62; p = 0.037), $Age^2$ (OR = 1.00, 95% CI 0.99, 1.00; p = 0.023) and the presence of COPD or HF (OR = 2.25, 95% CI1.30, 3.91; p = 0.004). The prediction accuracy of CRP and WBC did not reach the threshold of significance (p = 0.88 and p = 0.090).

Prognostic accuracy of MR-proADM in emergency departments

**Table 3.** Univariate regression analyses for predicting the three outcomes of interest: Acuity Increase, *Deterioration Event*, and *Length of Stay*. The analyses for the NEWS score as a predictor are shown in Table 4.

| | Beta | CI | Odds Ratio (CI) | p-value |
|---|---|---|---|---|
| **Acuity Increase: univariate logistic regressions (n = 292, e = 84)** | | | | |
| MR-proADM | 0.24 | -0.02, 0.48 | 1.27 (1.02, 1.62) | 0.037 |
| CRP | 0.003 | -0.0005, 0.0063 | 1.00 (1.00, 1.01) | 0.088 |
| WBC | 0.04 | -0.008, 0.094 | 1.05 (1.00, 1.10) | 0.09 |
| Gender | 0.14 | -0.38, 0.65 | 1.15 (0.69, 1.92) | 0.684 |
| Age | 0.1 | 0.019, 0.1925 | 1.11 (1.02, 1.21) | 0.023 |
| $Age^2$ | -0.0008 | -0.0016, -0.0001 | 1.00 (0.99, 1.00) | |
| Other Comorbidities | -0.32 | -0.96, 0.28 | 0.72 (0.38, 1.32) | 0.267 |
| COPD/HF* | 0.81 | 0.26, 1.36 | 2.25 (1.30, 3.91) | 0.004 |
| **Deterioration Event: univariate logistic regressions (n = 292, e2 = 32)** | | | | |
| MR-proADM | 0.37 | 0.11, 0.64 | 1.44 (1.12, 1.90) | 0.006 |
| CRP | 0.003 | -0.002, 0.01 | 1.00 (1.00, 1.01) | 0.255 |
| WBC | 0.02 | -0.05, 0.09 | 1.02 (0.95, 1.10) | 0.506 |
| Gender | 0.17 | -0.57, 0.92 | 1.19 (0.57, 2.50) | 0.648 |
| Age | 0.21 | 0.06, 0.40 | 1.23 (1.06, 1.49) | 0.013 |
| $Age^2$ | -0.002 | -0.003, -0.001 | 1.00 (1.00, 1.00) | |
| Other Comorbidities | 1.14 | 0.38, 1.90 | 3.14 (1.47, 6.69) | 0.003 |
| COPD/HF* | 0.67 | -0.14, 1.46 | 1.96 (0.87, 4.29) | 0.095 |
| **Length of stay: simple linear regressions (n = 292, e = 84, e2 = 32 )** | | | | |
| MR-proADM | 0.7 | 0.49, 0.92 | N/A | <0.0001 |
| CRP | 0.05 | -0.05, 0.15 | N/A | 0.368 |
| WBC | -0.06 | -0.38, 0.27 | N/A | 0.73 |
| Gender | 0.08 | -0.04, 0.20 | N/A | 0.18 |
| Age | 0.007 | 0.004, 0.010 | N/A | <0.0001 |
| Other Comorbidities | 0.18 | 0.05, 0.32 | N/A | 0.009 |
| COPD/HF* | 0.07 | -0.07, 0.21 | N/A | 0.318 |

**Key**: n = total number of cases; e = number of *Acuity Increases*; $e_2$ = number of *Deterioration Events*; CI = 95% confidence interval
* n = 282, e = 82, e2 = 29

**Table 4.** Multivariable regression analyses for the outcomes of interest: Acuity Increase, *Deterioration Event*, *Length of Stay* (Outcomes 1, 2, and 3 respectively) with NEWS comparator group. Predictor set *a.* includes only the NEWS score as a predictor; Predictor set *b.* includes MR-proADM and NEWS scores; Predictor set *c*. includes MR-proADM, NEWS scores, and other significant predictors and interactions.

| | | Beta | CI | Odds Ratio (CI) | p-value |
|---|---|---|---|---|---|
| ***Acuity Increase*: multivariate logistic regressions** | | | | | |
| **Predictor set a**<br>n = 292<br>e = 84<br>e2 = 32 | NEWS 3 | 0.06 | -0.55, 0.67 | 1.06 (0.57, 1.95) | 0.416 |
| | NEWS 4 | -0.48 | -1.29, 0.27 | 0.62 (0.27, 1.31) | |
| | NEWS 5 | 0.27 | -0.54, 1.04 | 1.31 (0.58, 2.84) | |
| **Predictor set *b*** | NEWS 3 | 0.03 | -0.59, 0.65 | 1.03 (0.56, 1.91) | 0.247 |
| n = 292<br>e = 84<br>e2 = 32 | NEWS 4 | -0.53 | -1.35, 0.23 | 0.59 (0.26, 1.26) | |
| | NEWS 5 | 0.18 | -0.63, 0.97 | 1.20 (0.53, 2.64) | |
| | MR-proADM | 0.24 | 0.02, 0.49 | 1.28 (1.02, 1.63) | 0.039 |
| **Predictor set c** | NEWS 3 | -0.11 | -0.76, 0.54 | 0.90 (0.47, 1.71) | 0.221 |
| n = 282<br>e = 82<br>e2 = 29 | NEWS 4 | -0.89 | -1.77, -0.08 | 0.41 (0.17, 0.93) | |
| | NEWS 5 | 0.09 | -0.77, 0.91 | 1.09 (0.46, 2.50) | |
| | MR-proADM | 0.41 | 0.13, 0.76 | 1.51 (1.14, 2.14) | 0.01 |
| | COPD/HF | 1.81 | 0.80, 2.85 | 6.08 (2.23, 17.35) | 0.001 |
| | MR-proADM*COPD/HF | -0.71 | -1.40, -0.10 | 0.49 (0.25, 0.91) | 0.03 |
| ***Deterioration Event*: multivariate logistic regressions** | | | | | |
| **Predictor set a** | NEWS 3 | 0.03 | -0.92, 0.94 | 1.03 (0.40, 2.55) | 0.512 |
| n = 292<br>e = 84<br>e2 = 32 | NEWS 4 | -0.37 | -1.68, 0.74 | 0.69 (0.19, 2.10) | |
| | NEWS 5 | 0.7 | -0.36, 1.70 | 2.02 (0.70, 5.50) | |
| **Predictor set *b*** | NEWS 3 | -0.01 | -0.97, 0.92 | 0.99 (0.38, 2.51) | 0.564 |
| *n* = 292<br>e = 84<br>e2 = 32 | NEWS 4 | -0.43 | -1.76, 0.70 | 0.65 (0.17, 2.02) | |
| | NEWS 5 | 0.6 | -0.49, 1.62 | 1.81 (0.61, 5.05) | |
| | MR-proADM | 0.36 | 0.10, 0.64 | 1.43 (1.11, 1.89) | 0.007 |
| **Predictor set c** | NEWS 3 | 0.16 | -0.83, 1.12 | 1.17 (0.44, 3.07) | 0.389 |
| n = 282<br>e = 82<br>e2 = 29 | NEWS 4 | -0.49 | -1.86, 0.69 | 0.62 (0.16, 2.00) | |
| | NEWS 5 | 0.69 | -0.44, 1.76 | 1.99 (0.64, 5.81) | |
| | MR-proADM | 0.32 | 0.02, 0.64 | 1.37 (1.02, 1.89) | 0.044 |

Prognostic accuracy of MR-proADM in emergency departments

| | | Beta | CI | Odds Ratio (CI) | p-value |
|---|---|---|---|---|---|
| | Other comorbidities | 0.94 | 0.10, 1.77 | 2.56 (1.10, 5.85) | 0.026 |
| | Age | 0.21 | 0.06, 0.41 | 1.23 (1.06, 1.50) | 0.011 |
| | Age$^2$ | -0.002 | -0.003, -0.001 | 1.00 (1.00, 1.00) | |
| ***Length of stay*: multiple linear regressions** | | | | | |
| **Predictor set a**<br>n = 292<br>e = 84<br>e2 = 32 | NEWS 3 | -0.07 | -0.21, 0.08 | N/A | 0.052 |
| | NEWS 4 | 0.07 | -0.10, 0.24 | N/A | |
| | NEWS 5 | 0.21 | 0.01, 0.40 | N/A | |
| **Predictor set *b*** | NEWS 3 | -0.1 | -0.24, 0.04 | N/A | 0.033 |
| *n = 292*<br>*e = 84*<br>*e2 = 32* | NEWS 4 | 0.05 | -0.11, 0.21 | N/A | |
| | NEWS 5 | 0.14 | -0.04, 0.32 | N/A | |
| | MR-proADM | 0.69 | 0.48, 0.91 | N/A | <0.0001 |
| **Predictor set c** | NEWS 3 | -0.12 | -0.25, 0.02 | N/A | 0.031 |
| n = 282<br>e = 82<br>e2 = 29 | NEWS 4 | 0.04 | -0.11, 0.20 | N/A | |
| | NEWS 5 | 0.14 | -0.04, 0.32 | N/A | |
| | MR-proADM | 0.55 | 0.31, 0.80 | N/A | <0.0001 |
| | Age | 0.004 | 0, 0.007 | N/A | 0.027 |

The predictive accuracy for *Acuity Increase* of NEWS on its own was limited (AUC 0.55, 95% CI 0.48, 0.62), but when MR-proADM was included as an additional predictor, the accuracy of the model increased substantially (AUC 0.61, 95% CI 0.54, 0.69; OR = 1.28, 95% CI 1.02, 1.63; p = 0.007) (Tables 4 and 5, Figure 2A), and was statistically significant (p = 0.033 for likelihood ratio, Table 5). When including MR-proADM with NEWS, the reclassification of patients was also significant, especially in the NRI score (NRI = 0.3, SE 0.1, p = 0.007; IDI = 0.017, Table 4).

The prediction accuracy of MR-proADM and the additional value it provides to the NEWS score was confirmed for *Deterioration Events* and *Length of Stay* (Figure 2C, Tables 4 and 5).

Prognostic accuracy of MR-proADM in emergency departments

**Table 5.** Model comparisons. Outcomes 1, 2, and 3 refer to Acuity Increase, *Deterioration Event*, and *Length of Stay* respectively. The predictors are: *Set a* NEWS score alone; *Set b* NEWS score and MR-proADM; *Set c* NEWS score, MR-proADM, and other significant predictors and interactions detailed in Table 3.

| | AIC | Deviance | AUC (CI) or $R^2$ for linear regression | LR, (df) p-value | NRI (se), p-value | IDI (se), p-value |
|---|---|---|---|---|---|---|
| *Acuity Increase*: logistic regressions | | | | | | |
| Outcome 1 - predictor set *a.* | 348 | 356 | 0.55 (0.48, 0.62) | | | |
| Outcome 1 - predictor set *b.* | 343 | 353 | 0.61 (0.54, 0.69) | 5 (1), 0.033 | 0.3 (0.1), 0.007 | 0.017 (0.009), 0.058 |
| Outcome 1 - predictor set *c.* | 317 | 331 | 0.69 (0.63, 0.76) | 14 (2), 0.001* | 0.4 (0.1), 0.0004* | 0.05 (0.01),0.0009* |
| *Deterioration Event*: logistic regressions | | | | | | |
| Outcome 2 - predictor set *a.* | 199 | 207 | 0.57 (0.47, 0.68) | | | |
| Outcome 2 - predictor set *b.* | 192 | 202 | 0.65 (0.54, 0.76) | 7 (1), 0.007 | 0.4 (0.2), 0.003 | 0.04 (0.02), 0.10 |
| Outcome 2 - predictor set *c.* | 177 | 193 | 0.73 (0.63, 0.84) | 15 (3), 0.0019* | 0.5 (0.2), 0.012* | 0.06 (0.02), 0.0004* |
| *Length of Stay*: linear regressions (LR) | | | | | | |
| Outcome 3 - predictor set *a.* | 77 | -381 | 0.03 | | | |
| Outcome 3 - predictor set *b.* | 68 | -417 | 0.14 | 9 (1), <0.001 | | |
| Outcome 3 - predictor set *c.* | 67 | -420 | 0.16 | 1 (1), 0.026 | | |

**Note**: AIC = Akaike information criterion; AUC = area under the receiver operating characteristic curve; CI = 95% confidence interval; LR = likelihood ratio; df = degrees of freedom; NRI = net reclassification index; se = standard error; IDI = integrated discrimination improvement. * Comparison is between predictor set b. and c. Since there was a mismatch between the cases for predictor set a. and b. (10 missing values in COPD/HF), in the model with predictors set b. the 10 cases missing in predictor set c. were dropped to allow the comparison.

## Effect on prediction accuracy when clinical information is added to the set of predictors

Ten additional patients with incomplete data were excluded from this analysis.

Multivariable modelling evaluated the predictive accuracy of MR-proADM when adjusted for the clinical factors in predictive set *c*: age, gender, CRP, WBC, presence of COPD or HF, presence of other comorbidities,

For *Acuity Increase,* COPD or HF comorbidity status and its interaction with MR-proADM level significantly improved the predictive accuracy of the model: AUC increased from 0.61 (95% CI 0.54, 0.69) to 0.69 (95% CI 0.63, 0.76). The increased risk of *Acuity Increase* with a unit increase in MR-proADM was 0.41 (95% CI 0.13, 0.76) with OR of 1.51 (95% CI1.14, 2.14; p = 0.010). The net reclassification index was significant (NRI = 0.4 (SE 0.1, p = 0.0004).

For *Deterioration Event*s, the presence of other comorbidities (excluding COPD and HF) and $Age^2$ increased the prediction accuracy of MR-proADM, (Table 4 and 5). The prediction accuracy of *Length of Stay* (Outcome 3) of MR-proADM is also increased including *Age* in the model (Table 4 and 5).

Because the means and standard errors of the coefficients estimated in the non-parametric bootstrapping analysis were all within 10-20% of the values evaluated by the models, the models' beta coefficients were not adjusted.

## Potential confounding effects

**Shorter term outcomes**: NEWS and MR-proADM had lower accuracy in predicting *Acuity Increase* within 24 and 12 hours from admission than in predicting *Acuity Increase* within 72 hours (Supplementary material, Tables 1 and 2).

**Interval between admission NEWS scoring and blood collection:** Because ward processes did not allow the times of scoring NEWS and collecting blood to be specified for research, we assessed for a confounding effect from variation in the timings, but found no evidence for it (Supplementary material, Table 3).

# Discussion

## Accuracy of prediction of deterioration by MR-proADM

This study shows that MR-proADM may be a clinically useful biomarker for predicting deterioration (i.e. *Acuity Increase*) within 72 hours from admission to hospital in patients with an admission NEWS score of 2 to 5. This contrasts with the performance of the NEWS score, assessed on admission, which did not predict deterioration within 72 hours, as might have been expected from previous evaluations [3-5 33]. This discrepancy with previous studies might be explained by differences in selection criteria for patients. Previous research included all patients admitted to ED, but our study selected patients with NEWS between 2 and 5, because a tool to predict deterioration would be most useful in this group.

For most of the observed *Acuity Increase* events, the reason for classification was an increase in the NEWS score. Because an increase in NEWS score reflects both measurement variation and physiological variation, additional exploratory analyses were carried out to assess the performance of MR-proADM with an operational definition of deterioration, *Deterioration Event*, designed to minimize measurement variation. NEWS on its own had low predictive accuracy for *Deterioration Event*s. However, MR-proADM level, and NEWS score together predicted *Deterioration Events* with an AUC of 0.65. Adding baseline patient characteristics (that were statistically selected) further increased the accuracy of the model (AUC = 0.73).

## Comorbidities and interactions with MR-proADM levels

MR-proADM levels in people with COPD and/or heart failure are chronically raised and are not predictive of deterioration. However, in other people whose MR-proADM levels are not chronically raised, high levels are predictive of *Acuity Increase* (Supplementary material, Figure 1). Including these comorbidities and their interaction with MR-proADM level increased the predictive accuracy of the logistic regression model.

## Limitations

This study included only patients who were admitted with a NEWS score between 2 and 5. The predictive accuracy of the MR-proADM would perhaps have been greater if more extreme cases had been included. However, patients with NEWS scores more than 5 are

known to be severely ill and to require close monitoring and/or management at higher levels of care.

## Interpretations and implications

The significant increase in predictive accuracy of the models when basic clinical information is added to the models suggests that value could be added to the NEWS score by using a clinical decision aid (CDA) that would have the NEWS score, MR-proADM level, age, and the presence of comorbidities as its inputs, and as its outputs, a risk score and advice on management decisions about the level of care and intensity of monitoring.

## Future research and development

As a growing number of NHS hospitals are implementing the NEWS score on their clinical information systems, it should be practical to develop a decision aid based on admission NEWS score, MR-proADM level, and possibly other clinical data. Other biomarkers may further improve prediction accuracy for deterioration, for example: lactate [3]; peroxiredoxin-4 (Prx4) and copeptin [22 34 35]; and soluble urokinase plasminogen activator receptor (suPAR)[36]. Developing CDAs with multiple biomarkers should increase the accuracy of prediction in ED and MAU where patients have many different conditions. The feasibility, cost-effectiveness, and acceptability of such a decision aid needs to be evaluated in further research.

A rapid point of care test could facilitate the assessment process and reduce delays in arranging optimal levels of care and intensity of monitoring.

# Footnotes

**Contributors:** JS, AJS, MP, and DS devised the study; RO, SG, and AJA managed the project; SG, and AJA performed the statistical analyses with advice from DS; all authors contributed to the final manuscript.

**Acknowledgements**: The authors would like to thank the patients who accepted to be part of the study, and the team of research nurses for the data collection: Laura Shewan, Louise Taylor, Janine Gradwell, Karen Martin, Katherine Cullen, Gerry Jones, Graham Soulsby, Carmen Scott, and Helen Reed.

The Diagnostic Evidence Co-operative (DEC) Newcastle is funded by the National Institute for Health Research (NIHR).

**Funding:** The study was funded by B·R·A·H·M·S GmbH

**Competing interests**: None declared.

**Ethics and governance approvals:** The study was approved by the Newcastle & North Tyneside Research Ethics Committee (15/NE/0120), and by the R&D Committee of the Newcastle upon Tyne Hospitals NHS foundation Trust (reference number 7495).

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data are available

## References

1. Royal College of Physicians of London. National early warning score (NEWS): standardising the assessment of acute-illness severity in the NHS—report of a working party. London, 2012.
2. McGinley A, Pearse RM. A national early warning score for acutely ill patients. *BMJ : British Medical Journal* 2012;345 doi: 10.1136/bmj.e5310
3. Jo S, Yoon J, Lee JB, et al. Predictive value of the National Early Warning Score–Lactate for mortality and the need for critical care among general emergency department patients. *Journal of Critical Care* 2016;36:60-68. doi: 10.1016/j.jcrc.2016.06.016 [published Online First: 20/06/29]
4. Smith GB, Prytherch DR, Meredith P, et al. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. *Resuscitation* 2013;84(4):465-70. doi: 10.1016/j.resuscitation.2012.12.016 [published Online First: 2013/01/04]
5. Smith ME, Chiovaro JC, O'Neil M, et al. Early warning system scores for clinical deterioration in hospitalized patients: a systematic review. *Annals of the American Thoracic Society* 2014;11(9):1454-65. doi: 10.1513/AnnalsATS.201403-102OC. [published Online First: 2014/10/09]
6. Teasdale GM. National early warning score (NEWS) is not suitable for all patients. *BMJ : British Medical Journal* 2012;345 doi: 10.1136/bmj.e5875
7. Eccles SR, Subbe C, Hancock D, et al. CREWS: Improving specificity whilst maintaining sensitivity of the National Early Warning Score in patients with chronic hypoxaemia. *Resuscitation* 2014;85(1):109-11. doi: 10.1016/j.resuscitation.2013.08.277 [published Online First: 20/09/2013]
8. Badriyah T, Briggs JS, Meredith P, et al. Decision-tree early warning score (DTEWS) validates the design of the National Early Warning Score (NEWS). *Resuscitation* 2014;85(3):418-23. doi: 10.1016/j.resuscitation.2013.12.011 [published Online First: 2013/12/19]

9. Abbott TE, Vaid N, Ip D, et al. A single-centre observational cohort study of admission National Early Warning Score (NEWS). *Resuscitation* 2015;92:89-93. doi: 10.1016/j.resuscitation.2015.04.020 [published Online First: 2015/05/03]

10. Corfield AR, Lees F, Zealley I, et al. Utility of a single early warning score in patients with sepsis in the emergency department. *Emergency medicine journal : EMJ* 2014;31(6):482-7. doi: 10.1136/emermed-2012-202186 [published Online First: 2013/03/09]

11. Bilben B, Grandal L, Søvik S. National Early Warning Score (NEWS) as an emergency department predictor of disease severity and 90-day survival in the acutely dyspneic patient – a prospective observational study. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine* 2016;24:80. doi: 10.1186/s13049-016-0273-9.

12. Valenzuela-Sanchez F, Valenzuela-Mendez B, Rodriguez-Gutierrez JF, et al. New role of biomarkers: mid-regional pro-adrenomedullin, the biomarker of organ failure. *Annals of translational medicine* 2016;4(17):329. doi: 10.21037/atm.2016.08.65 [published Online First: 2016/10/08]

13. Stokes NR, Dietz BW, Liang JJ. Cardiopulmonary laboratory biomarkers in the evaluation of acute dyspnea. *Open access emergency medicine : OAEM* 2016;8:35-45. doi: 10.2147/OAEM.S71446 [published Online First: 2016/06/17]

14. Sinning C, Zengin E, Zeller T, et al. Candidate biomarkers in heart failure with reduced and preserved ejection fraction. *Biomarkers : biochemical indicators of exposure, response, and susceptibility to chemicals* 2015;20(4):258-65. doi: 10.3109/1354750x.2015.1068856 [published Online First: 2015/08/21]

15. Kutz A, Hausfater P, Amin D, et al. The TRIAGE-ProADM Score for an Early Risk Stratification of Medical Patients in the Emergency Department - Development Based on a Multi-National, Prospective, Observational Study. *PloS one* 2016;11(12) doi: 10.1371/journal.pone.0168076 [published Online First: 2016/12/23]

16. Albrich WC, Dusemund F, Ruegger K, et al. Enhancement of CURB65 score with proadrenomedullin (CURB65-A) for outcome prediction in lower respiratory tract infections: derivation of a clinical algorithm. *BMC infectious diseases* 2011;11:112. doi: 10.1186/1471-2334-11-112 [published Online First: 2011/05/05]

17. Lopez J, Martinez A. Cell and molecular biology of the multifunctional peptide, adrenomedullin. *International review of cytology* 2002;221:1-92. [published Online First: 2002/11/29]

18. Serrano-Ponz M, Rodrigo-Gasque C, Siles E, et al. Temporal profiles of blood pressure, circulating nitric oxide, and adrenomedullin as predictors of clinical outcome in acute ischemic stroke patients. *Molecular medicine reports* 2016;13(5):3724-34. doi: 10.3892/mmr.2016.5001 [published Online First: 2016/03/18]

19. Zudaire E, Portal-Nunez S, Cuttitta F. The central role of adrenomedullin in host defense. *Journal of leukocyte biology* 2006;80(2):237-44. doi: 10.1189/jlb.0206123 [published Online First: 2006/06/12]

20. Zuur-Telgen MC, Brusse-Keizer MG, Vandervalk PD, et al. Stable state MR-proadrenomedullin level is a strong predictor for mortality in COPD patients. *Chest* 2013 doi: 10.1378/chest.13-1063 [published Online First: 2013/10/01]

21. Pedowska-Wloszek J, Kostrubiec M, Kurnicka K, et al. Midregional proadrenomedullin (MR-proADM) in the risk stratification of patients with acute pulmonary embolism. *Thrombosis research* 2013;132(5):506-10. doi: 10.1016/j.thromres.2013.08.008 [published Online First: 2013/08/19]

22. Nickel CH, Messmer AS, Geigy N, et al. Stress markers predict mortality in patients with nonspecific complaints presenting to the emergency department and may be a useful risk stratification tool to support disposition planning. *Academic emergency medicine : official journal of the Society for Academic Emergency Medicine* 2013;20(7):670-9. doi: 10.1111/acem.12172 [published Online First: 2013/07/19]

23. Landman GW, van Dijk PR, Drion I, et al. Midregional fragment of proadrenomedullin, new-onset albuminuria, and cardiovascular and all-cause mortality in patients with type 2 diabetes (ZODIAC-30). *Diabetes care* 2014;37(3):839-45. doi: 10.2337/dc13-1852 [published Online First: 2013/10/29]

24. Funke-Kaiser A, Mann K, Colquhoun D, et al. Midregional proadrenomedullin and its change predicts recurrent major coronary events and heart failure in stable coronary heart disease patients: the LIPID study. *International journal of cardiology* 2014;172(2):411-8. doi: 10.1016/j.ijcard.2014.01.064 [published Online First: 2014/01/24]

25. Gaggin HK, Januzzi JL, Jr. Biomarkers and diagnostics in heart failure. *Biochimica et biophysica acta* 2013;1832(12):2442-50. doi: 10.1016/j.bbadis.2012.12.014 [published Online First: 2013/01/09]

26. Vigue B, Leblanc PE, Moati F, et al. Mid-regional pro-adrenomedullin (MR-proADM), a marker of positive fluid balance in critically ill patients: results of the ENVOL study. *Critical care (London, England)* 2016;20(1):363. doi: 10.1186/s13054-016-1540-x [published Online First: 2016/11/09]

27. Tzikas S, Keller T, Ojeda FM, et al. MR-proANP and MR-proADM for risk stratification of patients with acute chest pain. *Heart (British Cardiac Society)* 2013;99(6):388-95. doi: 10.1136/heartjnl-2012-302956 [published Online First: 2012/12/04]

28. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology* 1996;49(12):1373-9. [published Online First: 1996/12/01]

29. Travaglino F, Russo V, De Berardinis B, et al. Thirty and ninety days mortality predictive value of admission and in-hospital procalcitonin and mid-regional pro-adrenomedullin testing in patients with dyspnea. Results from the VERyfing DYspnea trial. *The American journal of emergency medicine* 2014;32(4):334-41. doi: 10.1016/j.ajem.2013.12.045 [published Online First: 2014/01/03]

30. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria. R Foundation for Statistical Computing, 2011.

31. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., et al. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Statistics in medicine* 2008;27(2):157-72; discussion 207-12. doi: 10.1002/sim.2929 [published Online First: 2007/06/15]

32. Steyerberg EW, Vedder MM, Leening MJ, et al. Graphical assessment of incremental value of novel markers in prediction models: From statistical to decision analytical perspectives. *Biometrical journal Biometrische Zeitschrift* 2015;57(4):556-70. doi: 10.1002/bimj.201300260 [published Online First: 2014/07/18]

33. Alam N, Vegting IL, Houben E, et al. Exploring the performance of the National Early Warning Score (NEWS) in a European emergency department. *Resuscitation* 2015;90:111-5. doi: 10.1016/j.resuscitation.2015.02.011 [published Online First: 2015/03/06]

34. Nickel CH, Ruedinger J, Misch F, et al. Copeptin and peroxiredoxin-4 independently predict mortality in patients with nonspecific complaints presenting to the emergency department.

*Academic emergency medicine : official journal of the Society for Academic Emergency Medicine* 2011;18(8):851-9. doi: 10.1111/j.1553-2712.2011.01126.x [published Online First: 2011/08/17]

35. Iversen K, Gotze JP, Dalsgaard M, et al. Risk stratification in emergency patients by copeptin. *BMC medicine* 2014;12:80. doi: 10.1186/1741-7015-12-80 [published Online First: 2014/05/16]

36. Rasmussen LJ, Ladelund S, Haupt TH, et al. Soluble urokinase plasminogen activator receptor (suPAR) in acute care: a strong marker of disease presence and severity, readmission and mortality. A retrospective cohort study. *Emergency medicine journal : EMJ* 2016 doi: 10.1136/emermed-2015-205444 [published Online First: 2016/09/02]

# Figure legends

**Figure 1.** Patient recruitment process.

**Figure 2. Panel A.** Predictive accuracy for Acuity Increase; predictor set *a*: NEWS; predictors set *b*: NEWS, MR-proADM; predictor set c: NEWS, MR-proADM, COPD/HF, interaction between MR-proADM and COPD/HF. **Panel B.** Comparisons as for panel **A** but for predicting a *Deterioration Event*; predictor set *c:* NEWS score, MR-proADM level, Age[2], other comorbidities. **Panel C.** *Length of Stay* predicted by MR-proADM level.

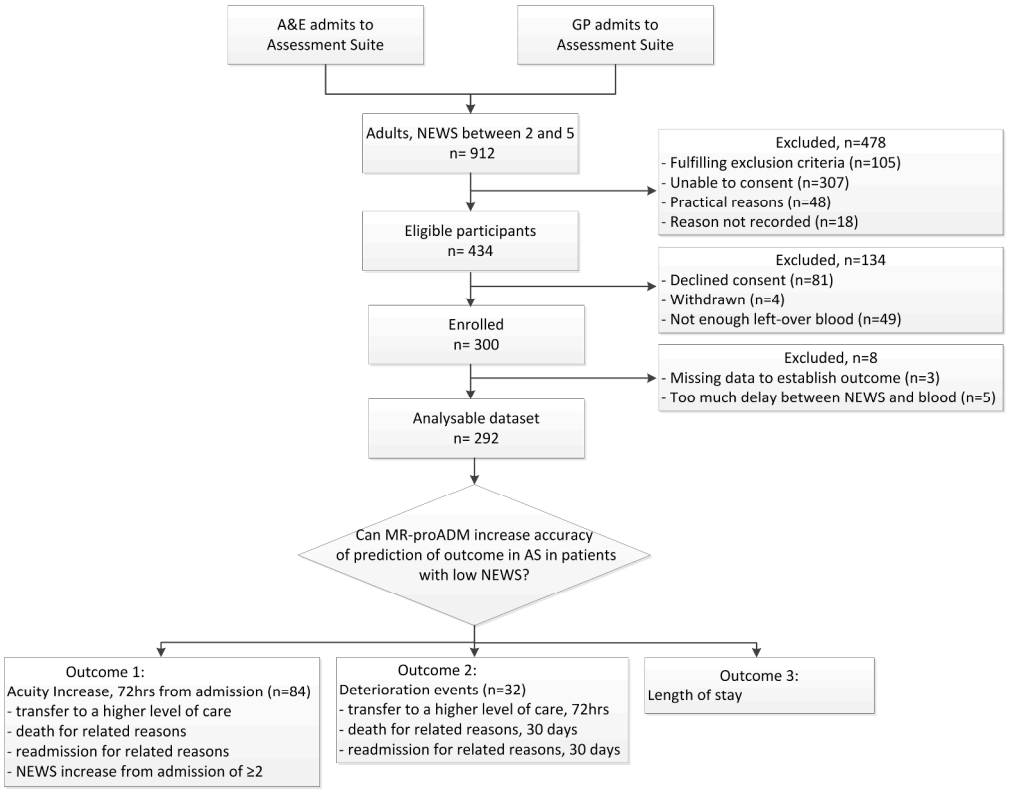



Figure 1. Patient recruitment process.

255x200mm (300 x 300 DPI)

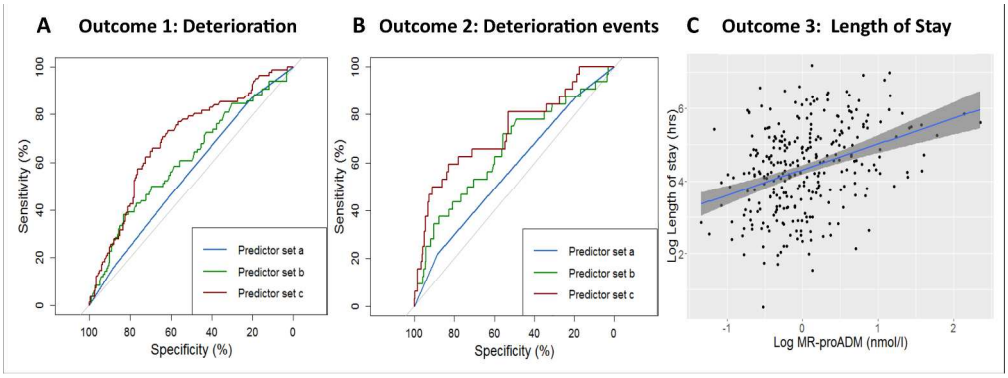



Figure 2. Panel A. Predictive accuracy for Acuity Increase; predictor set a: NEWS; predictors set b: NEWS, MR-proADM; predictor set c: NEWS, MR-proADM, COPD/HF, interaction between MR-proADM and COPD/HF. Panel B. Comparisons as for panel A but for predicting a Deterioration Event; predictor set c: NEWS score, MR-proADM level, Age2, other comorbidities. Panel C. Length of Stay predicted by MR-proADM level.

291x107mm (300 x 300 DPI)

Supplementary material for Graziadio et al, 2017

Supplementary material for *Can mid-regional pro-adrenomedullin (MR-proADM) increase the prognostic accuracy of NEWS in predicting deterioration in patients admitted to hospital with mild to moderately severe illness? A prospective single-centre observational study*. Graziadio et al, 2017

## Table of Contents

Supplementary material for Graziadio et al, 2017

# Additional information on Methods

## Visual data exploration and interaction between MR-proADM and COPD/HF

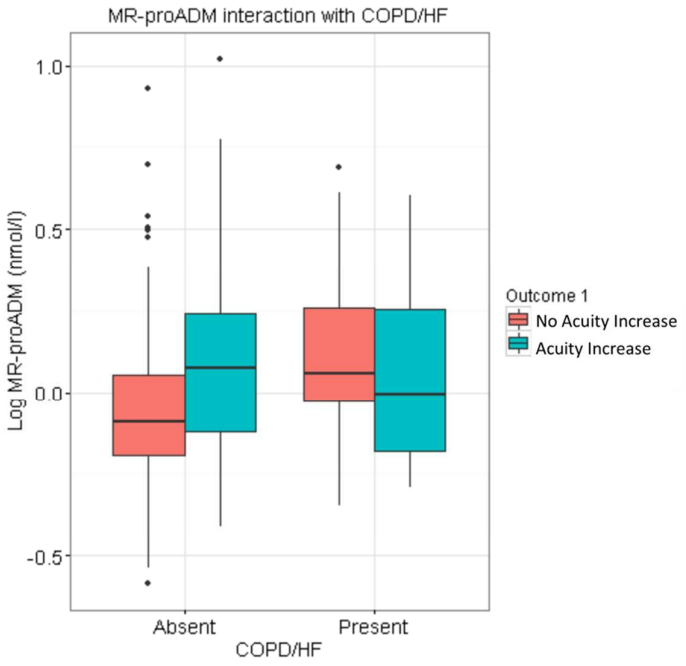
After formatting the datasets, all variables were graphed (bar-charts for categorical variables, and scatterplots/histograms for continuous variables) and visually checked for outliers and distributions that seemed potentially erroneous.

If outliers were identified, the cause(s) were investigated to understand whether they were due to human error or they were genuine data. Outliers were kept in the primary analysis. In a secondary sensitivity analysis, outliers were removed and the same analyses repeated to assess the impact on the results. If the coefficients of the predictors changed substantially, both models would be described. There was one genuine outliner patient with a very high level of MR-proADM compared to the population mean, but its exclusion made no difference to the results, and the subject was included in the final analysis.

The influence of potentially important factors on the ability of the MR-proADM to predict deterioration was explored graphically.

A significant interaction between MR-proADM and the presence of COPD/HF was discovered, and therefore included in the logistic regression (Outcome 1, predictor set c). The plot is shown in Supplementary Figure 1. This interaction showed that the MR-proADM level was increased in patients who deteriorated, but only if they did not have COPD or HF.

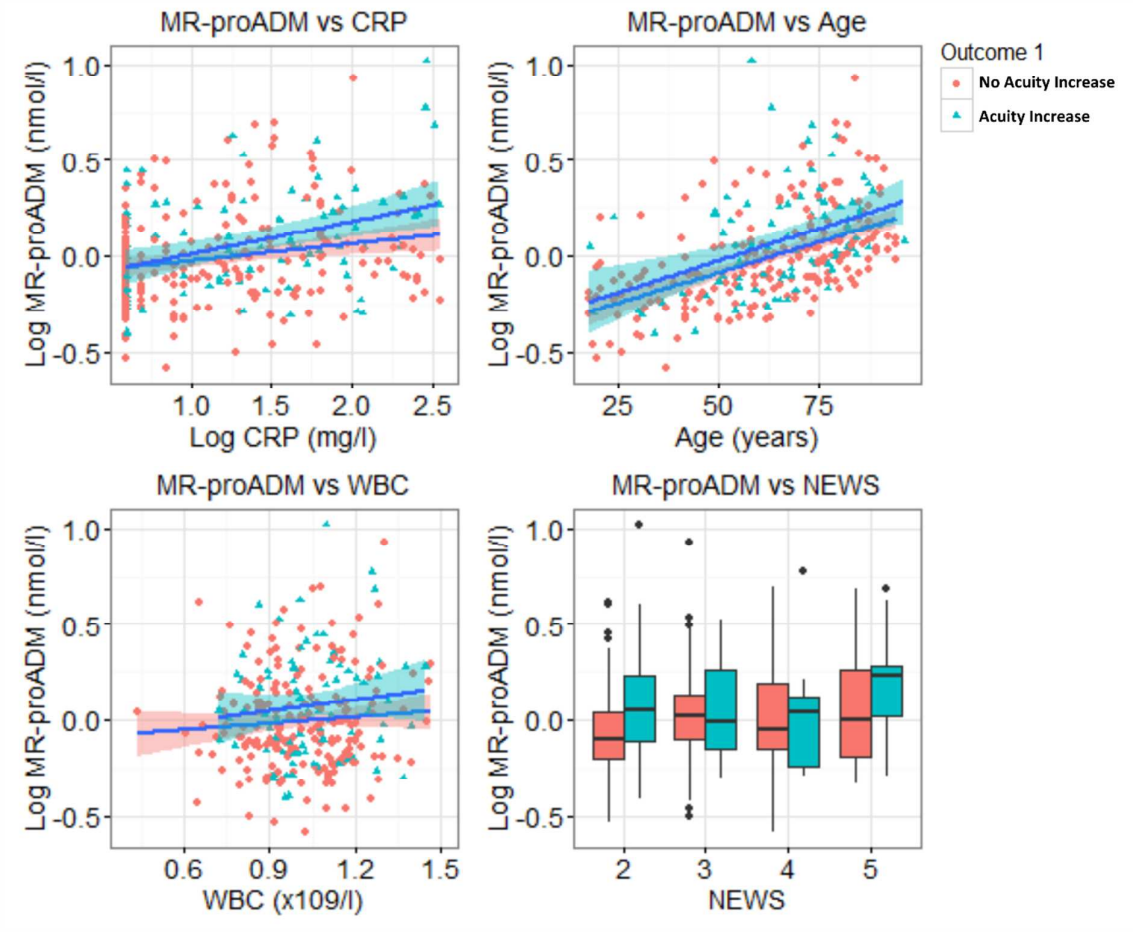
There was no suggestion that age; comorbidities: COPH and HF; other comorbidities; CRP; or WBC would improve the accuracy of prediction.




**Supplemental Figure 1**. Interaction between MR-proADM and COPD/HF.

Supplementary material for Graziadio et al, 2017

Correlations among biomarkers were also investigated through plotting to evaluate multicollinearity and added value of MR-proADM versus other biomarkers. Plots are shown in Supplemental Figure 2.




**Supplemental Figure 2.** Associations between MR-proADM and CRP, age, WBC, and NEWS.

## Analytical data exploration

Univariate logistic regressions were used to investigate whether the relationship between outcome variables (i.e. deterioration measures) and the input variables (NEWS and MR-proADM, age, comorbidities, gender, CRP, and WBC) were linear. If they were not linear, log transformation and squared transformation were applied. If the transformation substantially lowered the AIC, then the transformed variable was used in statistical analyses.

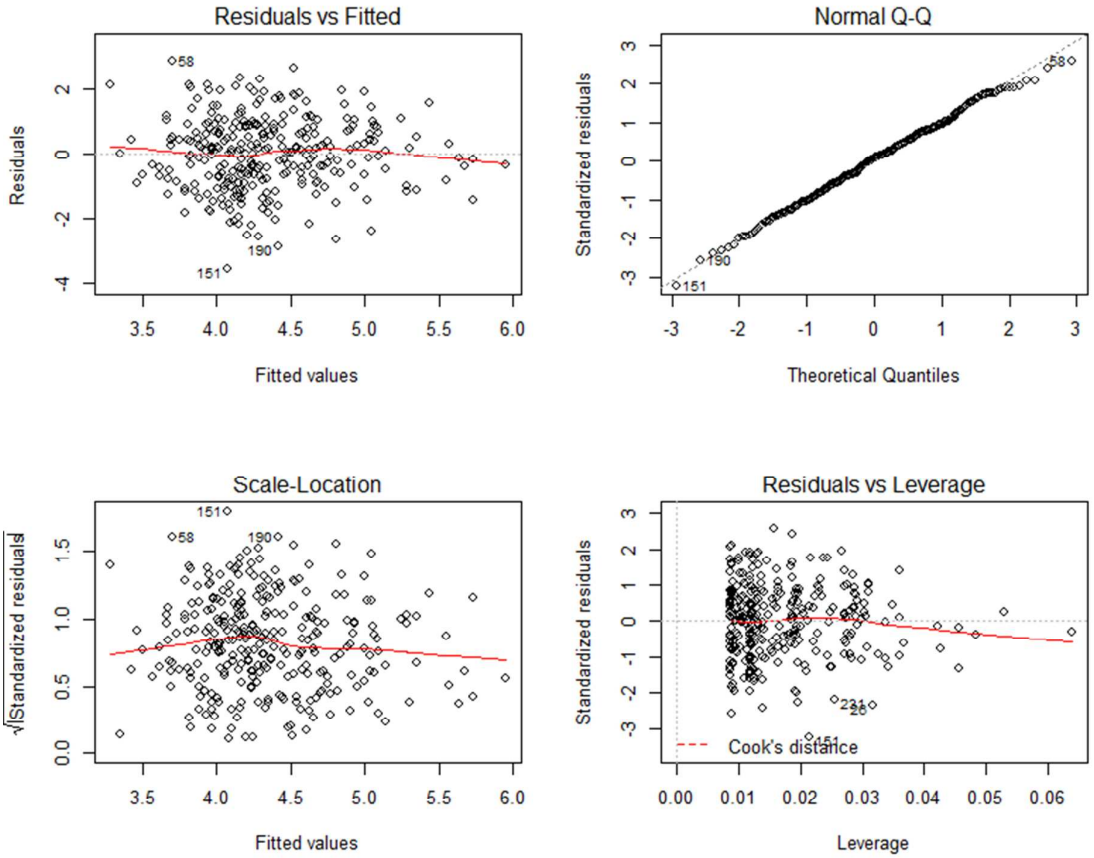
For categorical variables with multiple ordinal levels (i.e. NEWS score), the univariate analysis informed if it was appropriate to include the variable in the model as a continuous or categorical factor. If the coefficients in the univariate models increased linearly, then a linear relationship with the outcome could be assumed, and the variable was included in the model as continuous, otherwise the variable was treated as a categorical factor.

Supplementary material for Graziadio et al, 2017

Univariate analyses were also used to identify the variables that affected the outcome significantly. The variables that showed a probable relationship with the outcome variable ($p<0.1$) were included in the full model logistic regression.

# Additional information on Results

## Diagnostic plots for length of stay analysis

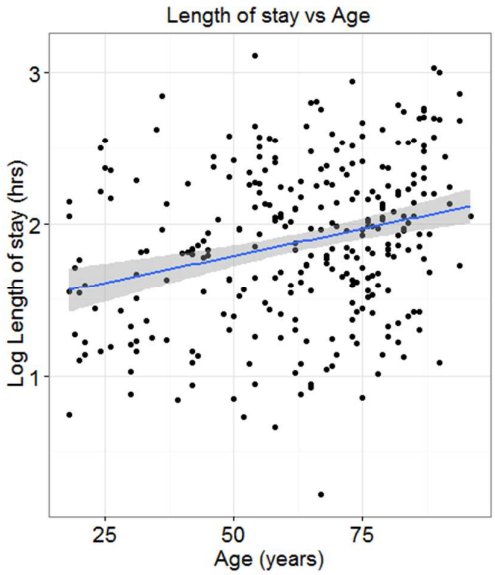



**Supplemental Figure 3.** Diagnostic plots for linear regression evaluating the prediction accuracy of MR-proADM for *Length of Stay*.

The diagnostics of the model showed no multicollinearity in the data since all the correlation coefficients among the independent variables were smaller than 0.5. No autocorrelation was found in the data, thus residuals are independent from each other: the Durbin-Watson test estimated d = 2.02 (p = 0.56). Evidence for homoscedasticity was provided graphically by the randomly scattered points and almost horizontal fitted lines in Supplemental Figure 3, (Residuals vs fitted plot). Analysis of Cook's distance showed that there were no influential points (d <4/51, Supplementary Figure 3).

In Supplemental Figure 4 the relationship between Length of Stay and Age is shown.

Supplementary material for Graziadio et al, 2017




**Supplemental Figure 4**. Relationship between *Length of Stay* and *Age*.

## Analyses of shorter term outcomes

The analyses found that NEWS and MR-proADM had much lower accuracy in predicting *Acuity Increase* at 24 and 12 hours from admission than in predicting *Acuity Increase* at 72 hours as apparent in Supplemental Tables 1 and 2.

**Supplemental Table 1.** Logistic regression for NEWS and MR-proADM predicting *Acuity Increase* within 24 hours. AIC = 326; AUC = 0.59.

| Covariate | Beta | CI | Odds ratio(CI) | P-value |
|---|---|---|---|---|
| Intercept | –1.42 | –1.70, –0.64 | 0.42 (0.27, 0.64) | NA |
| NEWS 3 | –0.006 | –0.64, 0.71 | 1.04 (0.53, 2.04) | 0.985 |
| NEWS 4 | –0.37 | –1.58, 0.20 | 0.52 (0.21, 1.22) | 0.368 |
| NEWS 5 | 0.21 | –0.66, 1.01 | 1.25 (0.52, 2.91) | 0.6244 |
| MR-proADM | 0.22 | –0.06, 0.44 | 1.20 (0.95, 1.55) | 0.0547 |

Supplementary material for Graziadio et al, 2017

**Supplemental Table 2.** Logistic regression for NEWS and MR-proADM predicting *Acuity Increase* within 12 hours. AIC = 266; AUC = 0.57.

| Covariate | Beta | CI | Odds ratio(CI) | P-value |
|---|---|---|---|---|
| **Intercept** | –1.83 | –2.47, –1.24 | 0.16 (0.08, 0.29) | NA |
| **NEWS 3** | 0.29 | –0.46, 1.04 | 1.34 (0.63, 2.85) | 0.442 |
| **NEWS 4** | –0.15 | –1.57, 0.76 | 0.86 (0.31, 2.14) | 0.756 |
| **NEWS 5** | 0.29 | –0.74, 1.23 | 1.33 (0.48, 3.42) | 0.564 |
| **MR-proADM** | 0.06 | –0.24, 0.31 | 1.06 (0.79, 1.36) | 0.656 |

## Analyses of time-lag effect between news assessment and blood collection for assessment of MR-proADM levels

Given the practicalities involved, it was not possible to stipulate the timings of taking the NEWS on admission and collecting the blood sample for MR-proADM testing. It was expected that difference in times would normally be less than 6 hours, but in 44 subjects the time difference was more than 6 hours.

To investigate the impact of time differences being greater than expected, another analysis was carried out excluding subjects for whom the difference was more than 6 hours (time-lag compliant dataset). The hypothesis was that, if the time difference was an important parameter for the predictive accuracy of MR-proADM level, model coefficients would be greater and confidence intervals narrower for the compliant model. This was not the case; results were similar in the full dataset with 292 subjects and in the compliant dataset with 248 subjects (Supplemental Table 3).

**Supplemental Table 3.** Logistic regression for NEWS and MR-proADM predicting *Acuity Increase* for the time-lag compliant dataset. AIC = 295; AUC = 0.60.

| Covariate | Beta | CI | Odds ratio(CI) | P-value |
|---|---|---|---|---|
| **Intercept** | –1.15 | –1.70, –0.64 | 0.42 (0.27, 0.64) | NA |
| **NEWS 3** | –0.04 | –0.64, 0.71 | 1.04 (0.53, 2.04) | 0.909 |
| **NEWS 4** | –0.65 | –1.58, 0.20 | 0.52 (0.21, 1.22) | 0.152 |
| **NEWS 5** | 0.22 | -0.66, 1.01 | 1.25 (0.52, 2.91) | 0.613 |
| **MR-proADM** | 0.19 | -0.06, 0.44 | 1.20 (0.95, 1.55) | 0.135 |

TRIPOD Checklist: Prediction Model Development

| Section/Topic | Item | Checklist Item | Page |
|---|---|---|---|
| **Title and abstract** | | | |
| Title | 1 | Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted. | P1 L1-3 |
| Abstract | 2 | Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions. | P2 L22-46 |
| **Introduction** | | | |
| Background and objectives | 3a | Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models. | P4-5 L52-84 |
| | 3b | Specify the objectives, including whether the study describes the development or validation of the model or both. | P5 L82-84 |
| **Methods** | | | |
| Source of data | 4a | Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable. | P5 L87 |
| | 4b | Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up. | P5 L87-91 |
| Participants | 5a | Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres. | P5 L87-97 |
| | 5b | Describe eligibility criteria for participants. | P5 L98-102 |
| | 5c | Give details of treatments received, if relevant. | N/A |
| Outcome | 6a | Clearly define the outcome that is predicted by the prediction model, including how and when assessed. | P6-7 L128-149 |
| | 6b | Report any actions to blind assessment of the outcome to be predicted. | P6 L123-126 |
| Predictors | 7a | Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured. | P4 L54-57 P4-5 L70-84 P5-6 L106-117 |
| | 7b | Report any actions to blind assessment of predictors for the outcome and other predictors. | P6 L123-126 |
| Sample size | 8 | Explain how the study size was arrived at. | P5 L92-97 |
| Missing data | 9 | Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method. | P8-9 L197-201 |
| Statistical analysis methods | 10a | Describe how predictors were handled in the analyses. | P7-8 L151-194 |
| | 10b | Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation. | P7-8 L151-194 P7 L153-155 |
| | 10d | Specify all measures used to assess model performance and, if relevant, to compare multiple models. | P7-8 L156-194 |
| Risk groups | 11 | Provide details on how risk groups were created, if done. | N/A |
| **Results** | | | |
| Participants | 13a | Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful. | P10-11 Table 1. |
| | 13b | Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome. | P10-11 Table 1. |
| Model development | 14a | Specify the number of participants and outcome events in each analysis. | P10-11 Table 1. |
| | 14b | If done, report the unadjusted association between each candidate predictor and outcome. | P13-15 Table3&4 |
| Model specification | 15a | Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point). | P13-15 Table3&4 |
| | 15b | Explain how to the use the prediction model. | P15 L244-253 |
| Model performance | 16 | Report performance measures (with CIs) for the prediction model. | P13-15 Table3&4 P15 L244-253 |
| **Discussion** | | | |
| Limitations | 18 | Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data). | P18-19 L306-316 |
| Interpretation | 19b | Give an overall interpretation of the results, considering objectives, limitations, and results from similar studies, and other relevant evidence. | P17-19 L278-323 |
| Implications | 20 | Discuss the potential clinical use of the model and implications for future research. | P19 L325-335 |
| **Other information** | | | |
| Supplementary | 21 | Provide information about the availability of supplementary resources, such as study | See Comments |

TRIPOD Checklist: Prediction Model Development

| | | | |
|---|---|---|---|
| information | | protocol, Web calculator, and data sets. | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study. | P20 L346 |

Comments:

1. **Item 21: Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.**

Supplementary material - with additional information on methods and results - is attached as separate document. Study protocol and data sets will be available in due course, new project website currently under construction.

BMJ Open

# Can mid-regional pro-adrenomedullin (MR-proADM) increase the prognostic accuracy of NEWS in predicting deterioration in patients admitted to hospital with mild to moderately severe illness? A prospective single-centre observational study.



| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2017-020337.R1 |
| Article Type: | Research |
| Date Submitted by the Author: | 31-May-2018 |
| Complete List of Authors: | Graziadio, Sara; Newcastle University, NIHR Newcastle In Vitro Diagnostics Co-operative ; Newcastle Upon Tyne Hospitals NHS Foundation Trust, NIHR Newcastle In Vitro Diagnostics Co-operative<br>Price, David; Newcastle Upon Tyne Hospitals NHS Foundation Trust, Department of Infectious Diseases<br>O'Leary, Rachel; Newcastle University, NIHR Newcastle In Vitro Diagnostics Co-operative<br>Stocken, Deborah; University of Leeds, Leeds Institute of Clinical Trials Research<br>Power, Michael; Newcastle Upon Tyne Hospitals NHS Foundation Trust, NIHR Newcastle In Vitro Diagnostics Co-operative<br>Allen, A; Newcastle University, NIHR Newcastle In Vitro Diagnostics Co-operative<br>Simpson, A John; Newcastle University Institute of Cellular Medicine, Institute of Cellular Medicine, Medical School; Newcastle Upon Tyne Hospitals NHS Foundation Trust, NIHR Newcastle In Vitro Diagnostics Co-operative |
| Primary Subject Heading: | Diagnostics |
| Secondary Subject Heading: | Evidence based practice |
| Keywords: | INTERNAL MEDICINE, ACCIDENT & EMERGENCY MEDICINE, Adult intensive & critical care < INTENSIVE & CRITICAL CARE |
| | |

Prognostic accuracy of MR-proADM in emergency departments

**Title:** Can mid-regional pro-adrenomedullin (MR-proADM) increase the prognostic accuracy of NEWS in predicting deterioration in patients admitted to hospital with mild to moderately severe illness? A prospective single-centre observational study.

**Authors:** Sara Graziadio[1], D. Ashley Price[2], Rachel A. O'Leary [1], Deborah D. Stocken[3], Michael Power[1], A. Joy Allen[1], A. John Simpson[1]

[1] National Institute for Health Research (NIHR) Newcastle In Vitro Diagnostics Cooperatives, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, NE1 4LP, UK

[2] Department of Infectious Diseases, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, NE1 4LP, UK

[3] Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, LS2 9JT, UK

**Corresponding author:** Dr Sara Graziadio, Newcastle University, NIHR Newcastle In Vitro Diagnostics Co-operative , Newcastle Upon Tyne, UK. Sara.Graziadio@ncl.ac.uk.

Word counts

Abstract: 276

Paper (including tables, figures, legends, and references): 5950

Prognostic accuracy of MR-proADM in emergency departments


## Abstract

**Objective** To assess the accuracy of NEWS score and MR-proADM blood level in predicting deterioration in mild to moderately ill people.

**Design** Prospective observational study

**Setting** The Medical Admissions Suite of the Royal Victoria Infirmary, Newcastle.

**Participants** 300 adults with NEWS score between 2 and 5 on admission. Exclusion criteria included receiving palliative care, or admitted for social reasons or self-harming. Patients were enrolled between September and December 2015, and followed-up for 30 days after discharge.

**Outcome measure** The primary outcome measure was the proportion of patients who, within 72 hours, had an *Acuity Increase*, defined as any combination of: an increase of at least 2 in the NEWS score; transferred to a higher-dependency bed or monitored area; and, for those discharged from hospital, re-admission for medical reasons; or death.

**Results** NEWS predicted *Acuity Increase* poorly: the area under the curve (AUC) was 0.55 (95% CI 0.48, 0.62) with univariate analysis. NEWS and MR-proADM together predicted *Acuity Increase* more accurately, increasing AUC to 0.61 (95% CI 0.54, 0.69). When the confounding effects of presence of chronic obstructive pulmonary disease or heart failure and interaction with MR-proADM were included, the prognostic accuracy further increased the AUC to 0.69 (95% CI 0.63, 0.76).

**Conclusions** MR-proADM improves the accuracy of prediction by NEWS of deterioration in patients admitted to hospital with a mild to moderately severe acute illness. As a growing number of NHS hospitals are implementing the NEWS score on their clinical information systems, further research should assess the practicalities and utility of developing a decision aid based on admission NEWS score, MR-proADM level, and possibly other clinical data and other biomarkers that could further improve prognostic accuracy.



## Keywords

Biochemistry, diagnosis, health services research

## Strengths and limitations of this study

- This is the first study to use rigorous statistical methods to assess the value added by MR-proADM to the admission NEWS score for predicting clinically important deterioration in mild to moderately ill patients
- Prognostic accuracy might have been greater had more severely ill patients been included, but the aim of this study was to predict deterioration in less severely ill patients who could benefit from closer observation.
- This was an observational study, and thus could not directly assess the utility of more accurate prediction of deterioration
- Initial evidence for MR-proADM appears promising and requires further validation for clinical utility

## Introduction

The National Early Warning Score (NEWS) is recommended for assessing severity of illness in patients presenting in primary or secondary NHS care and for surveillance of patients in hospital [1 2]. Six physiological parameters (which can be measured at the bedside) are scored: respiratory rate, oxygen saturation, temperature, systolic blood pressure, pulse rate, and level of consciousness. The scores are aggregated, and, if the patient requires oxygen, the total is increased. NEWS predicts death, cardiac arrest, and unplanned intensive care unit (ICU) admission within 24 hours [3-5]. However, NEWS does not identify all patients who turn out to be seriously ill [6-8], and there are also patients whose NEWS score is usually elevated and who do not require the level of observation that the NEWS tool would suggest. For example, people with chronic obstructive pulmonary disease (COPD) or chronic heart failure (HF) have higher baseline NEWS scores than those without these comorbidities. The prognostic accuracy of NEWS for patients presenting to the Emergency Department (ED) has been confirmed in a wide range of severity of illness [9 10], as has its reduced accuracy in people with COPD [11]. But, no previous studies of the prognostic accuracy of NEWS in the ED/Medical Admissions Unit (MAU) have focussed on patients admitted with mild to moderately severe illness. Since a clinically important proportion of these patients do deteriorate unexpectedly, improved risk stratification would be useful.

Mid-regional pro-adrenomedullin (MR-proADM) is one of several promising biomarkers for severe illness and deterioration [12-16].

MR-proADM is a precursor of adrenomedullin (ADM), a member of the calcitonin peptide family. ADM is widely expressed and has roles in vasodilation, immune modulation, and metabolic regulation. It is up-regulated in severe infections, inflammation, vasodilation, stimulation of diuresis, increased cardiac output, and stroke [17-19]. ADM has a short half-life, but MR-proADM is more stable and directly reflects ADM concentrations in blood. Both ADM and MR-proADM levels are strongly associated with risk of mortality, regardless of aetiology [20-26] . In people presenting with acute chest pain, MR-proADM has been reported to improve the Global Registry of Acute Coronary Events risk classification by 41% [27]. As with the NEWS score, people with COPD or chronic heart failure have higher baseline levels of MR-proADM.

The aim of this study was to assess whether the MR-proADM level used alongside the NEWS score would improve prediction of deterioration over NEWS score alone in patients admitted to the MAU with mild to moderately severe illness.

# Methods

## Patient and Public Involvement

Patients and the public were not specifically involved in the planning and execution of this study. However, the NIHR now requires that the research it supports includes active involvement and engagement with patients and the public.

## Study participants and study design

This was a prospective observational cohort study. Patients were enrolled between September and December 2015 at the Royal Victoria Infirmary, Newcastle, and followed-up for 30 days after discharge. If the patient died within the 30 days of follow up, this and cause of death were recorded. Adults admitted to the MAU were recruited for the study between 9am and 4pm on weekdays.

Sample size was determined as a pragmatic recruitment target for a three-month observational study. A recent unpublished audit conducted in the MAU at the Royal Victoria Infirmary found a deterioration rate of 20%. With 300 patients and complete data collection, 60 events would be anticipated. With this number of events, a multivariable prediction model could include up to six independent predictors. This is based on a widely accepted rule of thumb that models with fewer than ten events per predictor tend to be over-fitted [28]. However, recent research suggests that the "ten events per variable" rule of thumb may be optimistic [29]. Because the aim of this study was to assess if further research would be indicated, even if the rule of thumb is optimistic, 60 is considered an acceptable number of events.

Patients were considered eligible for inclusion in the study if their NEWS score on admission was at least 2 and not greater than 5, and all NEWS parameters were recorded. Patients were excluded from the study if they were receiving palliative care, were admitted for social reasons only, or were self-harming, or overdosing with drugs or other substances.

All participants provided written informed consent, and the study was approved by the Newcastle & North Tyneside Research Ethics Committee (15/NE/0120).

## Recorded data

Demographic and admission data included: gender, year of birth, reason for admissions, diagnosis on discharge, and the presence of comorbidities in which baseline MR-proADM levels are chronically raised: COPD with hypoxia ($PaO_2$ <10 kPa) [7]; HF [30]; acute brain injury [6]; acute coronary syndrome [27]; acute venous thromboembolism [21]; high International Normalized Ratio (INR>2); acute kidney injury; electrolyte disturbances ($Na^+$ <130 or >150 mmol/L; $K^+$ <3.0 or >5.5 mmol/L); hyperglycaemia in type 1 diabetes (random glucose >10 mmol/L).

The NEWS score was assessed at on admission and over the next 72 hours, and the scores and assessment times recorded. The 7 clinical parameters used to determine the NEWS score were recorded for the baseline (admission) assessment only. Baseline NEWS scores were used to determine eligibility for this study. Subsequent NEWS scores were used in the analyses to identify deterioration.

Blood samples were taken at hospital admission for assessment of MR-proADM, C-Reactive Protein (CRP) and white blood count (WBC).

## Laboratory tests

Plasma was obtained from blood samples (collected in ethylenediaminetetraacetic acid, EDTA) that were no longer clinically required. Plasma was stored in aliquots at –80º C.

MR-proADM was assayed in the on-site Blood Sciences Laboratory using the B R A H M S Kryptor system according to the manufacturer's instructions.

Blood samples were analysed in batches by personnel blinded with regard to the condition and NEWS score of the patient. Nurses who assessed the NEWS score and healthcare professionals managing patients in the MAU were blinded to MR-proADM results.

## Outcome measures

**Outcome measure 1: *Acuity Increase***. A patient was classified as having an *Acuity Increase* if one or more of the following occurred within 72 hours from admission:

1. transfer to a higher level of care (ICU or high dependency unit)
2. readmission to hospital for reasons related to the initial admission
3. death for reasons related to the initial admission
4. NEWS score increased by at least two compared to the admission score

**Outcome measure 2: *Deterioration Event***. For most of the observed *Acuity Increase* cases the reason for classification was an increase in the NEWS score (Table 1). An increase in NEWS score reflects both measurement variation and physiological variation, so additional exploratory analyses were carried out to assess the performance of MR-proADM in predicting deterioration. *Deterioration Events* were classified as the occurrence of one or more of the following:

1. transfer to higher level of care within 72 hours from admission;
2. death (for reasons related to the admission) within 30 days;
3. re-admission to hospital (for the same reason as the previous admission) within 30 days from first admission.

Classification based on this definition is unlikely to be subject to clinically important measurement variation. This analysis, therefore, should optimise the prognostic accuracy for events which are both clinically and economically important.

**Outcome measure 3: *Length of Stay***. *Length of Stay* was defined as the duration (in days) from admission to discharge or death.

## Statistical analysis

All data analyses were performed using the R language version 3.2.0 [31], with the support of RStudio, version 0.99.896 (RStudio, Inc). The following R packages were used: ggplot2, pROC, psych, PredictABEL, Hmisc, rms.

Logistic regression models were compared for their accuracy in predicting deterioration outcome measures as pre-specified in an analysis plan. Analyses are presented as unadjusted parameter estimates of risk (odds ratio (OR), with confidence intervals (CI)) and estimates adjusted for identified clinical confounding factors. The aims of the multivariable analyses were twofold: first, to estimate the effect size and significance adjusted for other identified influential predictors and interactions; second, to investigate whether the addition of other predictors improved the goodness of fit and accuracy of prediction.

Only complete cases were analysed since missingness was minimal: 10 records without data on co-morbidities (details in footnote in Table 1).

For each measure of deterioration (*Acuity Increase, Deterioration Event*, and *Length of Stay*), logistic regression models were compared for the following sets of predictor variables:

*Predictor set a.* Comparator (base case): NEWS score on admission

*Predictor set b.* Primary analysis: NEWS score, MR-proADM

*Predictor set c.* Secondary analyses: NEWS score and MR-proADM always included. Age, gender, CRP, WBC, presence of COPD or HF, presence of other comorbidities, and interactions between predictors when appropriate.

Predictors (and the underlying assumption of linearity of their relationship with the outcome of interest) were initially investigated through univariate analyses based on simple log and quadratic functions. Transformations were applied if they improved the goodness of fit as assessed by the Akaike information criterion (AIC), and were retained in the multivariable setting. NEWS was treated as an ordinal variable. We have assessed interactions through visual data exploration and acknowledge this is underpowered. Subsequently, for the multivariable regression the set of predictors was assessed for independence through backward elimination, based on changes in AIC.

Secondary outcome of *Length of Stay* followed a similar analysis plan using multiple linear regressions based on a transformed outcome to address non-normality. Dependent and exploratory variables were log-transformed if not normally distributed. Normality was assessed by visualizing the data. More details of the methods used are reported in the Supplemental Data.

Goodness of fit of logistic regression models was assessed with the C-statistic (which is the area under the ROC curve and a measure of discrimination) and is presented with 95% confidence intervals (CI). To assess the value added by including the MR-proADM level with the NEWS score in predicting deterioration, continuous net reclassification improvement (NRI) and integrated discrimination improvement (IDI) were calculated[32 33].

For internal validation of the statistical models the C-statistic was evaluated after correcting for optimistic predictions through bootstrapping with 10,000 resamples.

# Results

## Study enrolment

The process of recruitment and enrolment of patients for the study is shown in Figure 1. The study recruited 300 patients, and 292 were included in the analysis. Five patients were excluded because the blood samples for MR-proADM were taken more than 12 hours from baseline NEWS assessment; 3 patients were excluded from the primary outcome due to missing follow up NEWS scores.

## Patient characteristics

Patient demographics and mean biomarker levels for each covariate are reported in Table 1. The cohort was evenly divided in gender and had a mean age of 63 years and mean NEWS on admission of 3, with the majority of patients having NEWS score of 2. COPD or HF were present in 28%, and 25% had other comorbidities.

Prognostic accuracy of MR-proADM in emergency departments

**Table 1**. Characteristics of the study population, classified by Outcome 1 (*Acuity Increase*), Outcome 2 (*Deterioration Event*) and All patients. Data are presented as number (no) and percentages (%) for counts, or mean and (standard deviation, SD) for continuous normally distributed data, or [25th; 50th; 75th percentile] for continuous non-normally distributed data.

| | **Outcome 1: *Acuity Increase*** | | **Outcome 2: *Deterioration Event*** | | **All patients (n = 292)** |
|---|---|---|---|---|---|
| | **Present (e = 84)** | **Absent** | **Present (e2 = 32)** | **Absent** | |
| Age (mean years, SD) | 65 (17) | 62 (21) | 63 (14) | 63 (20) | 63 (20) |
| Gender (no. females, %) | 41 (49%) | 107 (51%) | 15 (47%) | 133 (51%) | 148 (51%) |
| NEWS = 2 (no., %) | 34 (40%) | 82 (40%) | 12 (38%) | 104 (40%) | 116 (40%) |
| NEWS = 3 (no., %) | 26 (31%) | 59 (28%) | 9 (28%) | 76 (29%) | 85 (29%) |
| NEWS = 4 (no., %) | 11 (13%) | 43 (21%) | 4 (13%) | 50 (19%) | 54 (18%) |
| NEWS = 5 (no., %) | 13 (15%) | 24 (12%) | 7 (22%) | 30 (12%) | 37 (13%) |
| MR-proADM (mean nmol/l, SD) | 1.50 (1.4)<br>[0.72, 1.12, 1.79] | 1.19 (0.9)<br>[0.68, 0.93, 1.28] | 1.89 (2.0)<br>[0.93, 1.13, 1.95] | 1.20 (0.9)<br>[0.68, 0.93, 1.39] | 1.28 (1.1)<br>[0.68, 0.97, 1.48] |
| CRP (mg/l) | 59 (79)<br>[5, 22, 80] | 42 (70)<br>[4, 13, 41] | 61 (90)<br>[7, 23, 67] | 45 (71)<br>[4, 16, 51] | 47 (73)<br>[4, 17, 54] |
| WBC (x$10^9$/l) | 12 (5)<br>[9, 10, 14] | 11 (5)<br>[8, 10, 14] | 12 (4)<br>[9, 12, 15] | 11 (5)<br>[8, 10, 14] | 11 (5)<br>[8, 10, 14] |
| COPD/HF (no, %)* | 33 (39%) | 46 (22%) | 12 (38%) | 67 (26%) | 79 (28%) |
| Other comorbidities (no., %) | 17 (20%) | 55 (26%) | 15 (47%) | 57 (22%) | 72 (25%) |
| *Length of Stay* (hrs) | 168 (196)<br>[63, 110, 194] | 137 (176)<br>[26, 68, 176] | 173 (172)<br>[59, 106, 259] | 143 (172)<br>[33, 72, 176] | 146 (182)<br>[35, 77, 182] |
| Length of Stay in MAU (hrs) | 31 (19)<br>[17, 25, 43] | 24 (16)<br>[13, 21, 30] | 27 (17)<br>[18, 23, 35] | 26 (17)<br>[15, 22, 31] | 26 (17)<br>[15, 22, 31] |
| Monitored beds (no, %) | 31 (37%) | 58 (27%) | 11 (34%) | 78 (30%) | 89 (30%) |

Prognostic accuracy of MR-proADM in emergency departments

| | Outcome 1: *Acuity Increase* | | Outcome 2: *Deterioration Event* | | All patients (n = 292) |
|---|---|---|---|---|---|
| | Present (e = 84) | Absent | Present (e2 = 32) | Absent | |
| Deterioration time (hrs) | 15 (13) [5, 9, 21] | N/A | 170 (226) [19, 33, 301] | N/A | |
| * for COPD: e = number with *Acuity Increase* = 82; e2 = number with *Deterioration Event = 29;* n = total number of patients = 282 | | | | | |

218

**Table 2**. Criteria met by patients classified with an *Acuity Increase* or *Deterioration Event.*

| Criterion for deterioration | *Acuity Increase* (e = 84) | *Deterioration Event* (e2 = 32) |
|---|---|---|
| NEWS (no, %) | 81 (96.4%) | N/A |
| ICU transfer (no, %) | 1 (1.2%) | 4 (12.5%) |
| Death (no, %) | 0 (0%) | 6 (18.8%) |
| Readmission (no, %) | 2 (2.4%) | 22 (68.7%) |

The study population was homogenous across *Acuity Increase* and No *Acuity Increase* outcomes in terms of gender, age, and NEWS on admission. Table 2 shows the frequencies of criteria determining *Acuity Increase* and *Deterioration Event*. Notably, around 95% of *Acuity Increases* were the result of an increase in NEWS score, while readmission was the reason for around 70% of *Deterioration Event*s.

Patients who experienced *Acuity Increase* had higher MR-proADM and CRP levels at admission, and longer *Length of Stay* in the hospital and in the MAU.

The prevalence of *Acuity Increase* was 29% (somewhat higher than the anticipated 20%). The prevalence of *Deterioration Event*s was 11%. The numbers of events provided sufficient statistical power to assess statistical significance for the primary outcome, *Acuity Increase*, but not for the secondary outcome, *Deterioration Event*.

## Accuracy of MR-proADM for predicting *Acuity Increase*

Potentially useful predictors with univariate analysis of *Acuity Increase* were MR-proADM (OR = 1.27, 95% CI 1.02, 1.62; p = 0.037), $Age^2$ (OR = 1.00, 95% CI 0.99, 1.00; p = 0.023) and the presence of COPD or HF (OR = 2.25, 95% CI1.30, 3.91; p = 0.004; Supplementary Figure s1). The prognostic accuracy of CRP, WBC and NEWS did not reach the threshold of significance (p = 0.88, p = 0.090, Table 3, and p=0.416, Table 4, respectively).

**Table 3.** Univariate regression analyses for predicting the three outcomes of interest: *Acuity Increase*, *Deterioration Event*, and *Length of Stay*. Analyses for the NEWS score as a predictor are shown in Table 4.

| | Beta | CI | Odds Ratio (CI) | p-value |
|---|---|---|---|---|
| ***Acuity Increase*: univariate logistic regressions (n = 292, e = 84)** | | | | |
| MR-proADM | 0.24 | -0.02, 0.48 | 1.27 (1.02, 1.62) | 0.037 |
| CRP | 0.003 | -0.0005, 0.0063 | 1.00 (1.00, 1.01) | 0.088 |
| WBC | 0.04 | -0.008, 0.094 | 1.05 (1.00, 1.10) | 0.09 |
| Gender | 0.14 | -0.38, 0.65 | 1.15 (0.69, 1.92) | 0.684 |
| Age | 0.1 | 0.019, 0.1925 | 1.11 (1.02, 1.21) | 0.023 |
| $Age^2$ | -0.0008 | -0.0016, -0.0001 | 1.00 (0.99, 1.00) | |
| Other Comorbidities | -0.32 | -0.96, 0.28 | 0.72 (0.38, 1.32) | 0.267 |
| COPD/HF* | 0.81 | 0.26, 1.36 | 2.25 (1.30, 3.91) | 0.004 |
| ***Deterioration Event*: univariate logistic regressions (n = 292, e2 = 32)** | | | | |
| MR-proADM | 0.37 | 0.11, 0.64 | 1.44 (1.12, 1.90) | 0.006 |
| CRP | 0.003 | -0.002, 0.01 | 1.00 (1.00, 1.01) | 0.255 |
| WBC | 0.02 | -0.05, 0.09 | 1.02 (0.95, 1.10) | 0.506 |
| Gender | 0.17 | -0.57, 0.92 | 1.19 (0.57, 2.50) | 0.648 |
| Age | 0.21 | 0.06, 0.40 | 1.23 (1.06, 1.49) | 0.013 |
| $Age^2$ | -0.002 | -0.003, -0.001 | 1.00 (1.00, 1.00) | |
| Other Comorbidities | 1.14 | 0.38, 1.90 | 3.14 (1.47, 6.69) | 0.003 |
| COPD/HF* | 0.67 | -0.14, 1.46 | 1.96 (0.87, 4.29) | 0.095 |
| ***Length of Stay*: simple linear regressions (n = 292, e = 84, e2 = 32 )** | | | | |
| MR-proADM | 0.7 | 0.49, 0.92 | N/A | <0.0001 |
| CRP | 0.05 | -0.05, 0.15 | N/A | 0.368 |
| WBC | -0.06 | -0.38, 0.27 | N/A | 0.73 |
| Gender | 0.08 | -0.04, 0.20 | N/A | 0.18 |
| Age | 0.007 | 0.004, 0.010 | N/A | <0.0001 |
| Other Comorbidities | 0.18 | 0.05, 0.32 | N/A | 0.009 |
| COPD/HF* | 0.07 | -0.07, 0.21 | N/A | 0.318 |

**Key**: n = total number of cases; e = number of *Acuity Increases*; $e_2$ = number of *Deterioration Events*; CI = 95% confidence interval
* n = 282, e = 82, e2 = 29

**Table 4.** Multivariable regression analyses for the outcomes of interest: *Acuity Increase*, *Deterioration Event*, *Length of Stay* (Outcomes 1, 2, and 3 respectively) with NEWS comparator group. Predictor set *a.* includes only the NEWS score as a predictor; Predictor set *b.* includes MR-proADM and NEWS scores; Predictor set *c*. includes MR-proADM, NEWS scores, and other significant predictors and interactions.

| | | Beta | CI | Odds Ratio (CI) | p-value |
|---|---|---|---|---|---|
| ***Acuity Increase*: multivariate logistic regressions** | | | | | |
| **Predictor set a**<br>n = 292<br>e = 84<br>e2 = 32 | NEWS 3 | 0.06 | -0.55, 0.67 | 1.06 (0.57, 1.95) | 0.416 |
| | NEWS 4 | -0.48 | -1.29, 0.27 | 0.62 (0.27, 1.31) | |
| | NEWS 5 | 0.27 | -0.54, 1.04 | 1.31 (0.58, 2.84) | |
| **Predictor set *b*** | NEWS 3 | 0.03 | -0.59, 0.65 | 1.03 (0.56, 1.91) | 0.247 |
| n = 292<br>e = 84<br>e2 = 32 | NEWS 4 | -0.53 | -1.35, 0.23 | 0.59 (0.26, 1.26) | |
| | NEWS 5 | 0.18 | -0.63, 0.97 | 1.20 (0.53, 2.64) | |
| | MR-proADM | 0.24 | 0.02, 0.49 | 1.28 (1.02, 1.63) | 0.039 |
| **Predictor set c** | NEWS 3 | -0.11 | -0.76, 0.54 | 0.90 (0.47, 1.71) | 0.221 |
| n = 282<br>e = 82<br>e2 = 29 | NEWS 4 | -0.89 | -1.77, -0.08 | 0.41 (0.17, 0.93) | |
| | NEWS 5 | 0.09 | -0.77, 0.91 | 1.09 (0.46, 2.50) | |
| | MR-proADM | 0.41 | 0.13, 0.76 | 1.51 (1.14, 2.14) | 0.01 |
| | COPD/HF | 1.81 | 0.80, 2.85 | 6.08 (2.23, 17.35) | 0.001 |
| | MR-proADM*COPD/HF | -0.71 | -1.40, -0.10 | 0.49 (0.25, 0.91) | 0.03 |
| ***Deterioration Event*: multivariate logistic regressions** | | | | | |
| **Predictor set a** | NEWS 3 | 0.03 | -0.92, 0.94 | 1.03 (0.40, 2.55) | 0.512 |
| n = 292<br>e = 84<br>e2 = 32 | NEWS 4 | -0.37 | -1.68, 0.74 | 0.69 (0.19, 2.10) | |
| | NEWS 5 | 0.7 | -0.36, 1.70 | 2.02 (0.70, 5.50) | |
| **Predictor set *b*** | NEWS 3 | -0.01 | -0.97, 0.92 | 0.99 (0.38, 2.51) | 0.564 |
| *n* = 292<br>e = 84<br>e2 = 32 | NEWS 4 | -0.43 | -1.76, 0.70 | 0.65 (0.17, 2.02) | |
| | NEWS 5 | 0.6 | -0.49, 1.62 | 1.81 (0.61, 5.05) | |
| | MR-proADM | 0.36 | 0.10, 0.64 | 1.43 (1.11, 1.89) | 0.007 |
| **Predictor set c** | NEWS 3 | 0.16 | -0.83, 1.12 | 1.17 (0.44, 3.07) | 0.389 |
| n = 282<br>e = 82<br>e2 = 29 | NEWS 4 | -0.49 | -1.86, 0.69 | 0.62 (0.16, 2.00) | |
| | NEWS 5 | 0.69 | -0.44, 1.76 | 1.99 (0.64, 5.81) | |
| | MR-proADM | 0.32 | 0.02, 0.64 | 1.37 (1.02, 1.89) | 0.044 |

| | | Beta | CI | Odds Ratio (CI) | p-value |
|---|---|---|---|---|---|
| | Other comorbidities | 0.94 | 0.10, 1.77 | 2.56 (1.10, 5.85) | 0.026 |
| | Age | 0.21 | 0.06, 0.41 | 1.23 (1.06, 1.50) | 0.011 |
| | Age$^2$ | -0.002 | -0.003, -0.001 | 1.00 (1.00, 1.00) | |
| ***Length of Stay*: multiple linear regressions** | | | | | |
| **Predictor set a**<br>n = 292<br>e = 84<br>e2 = 32 | NEWS 3 | -0.07 | -0.21, 0.08 | N/A | 0.052 |
| | NEWS 4 | 0.07 | -0.10, 0.24 | N/A | |
| | NEWS 5 | 0.21 | 0.01, 0.40 | N/A | |
| **Predictor set *b*** | NEWS 3 | -0.1 | -0.24, 0.04 | N/A | 0.033 |
| *n = 292*<br>*e = 84*<br>*e2 = 32* | NEWS 4 | 0.05 | -0.11, 0.21 | N/A | |
| | NEWS 5 | 0.14 | -0.04, 0.32 | N/A | |
| | MR-proADM | 0.69 | 0.48, 0.91 | N/A | <0.0001 |
| **Predictor set c** | NEWS 3 | -0.12 | -0.25, 0.02 | N/A | 0.031 |
| n = 282<br>e = 82<br>e2 = 29 | NEWS 4 | 0.04 | -0.11, 0.20 | N/A | |
| | NEWS 5 | 0.14 | -0.04, 0.32 | N/A | |
| | MR-proADM | 0.55 | 0.31, 0.80 | N/A | <0.0001 |
| | Age | 0.004 | 0, 0.007 | N/A | 0.027 |

The prognostic accuracy for *Acuity Increase* of NEWS on its own was not significant and limited (AUC 0.55, 95% CI 0.48, 0.62), but when MR-proADM was included as an additional predictor, the accuracy of the model increased (AUC 0.61, 95% CI 0.54, 0.69; OR = 1.28, 95% CI 1.02, 1.63; p = 0.007) (Tables 4 and 5, Figure 2 panel A), and was statistically significant (p = 0.033 for likelihood ratio, Table 5). When including MR-proADM with NEWS, the reclassification of patients was also significant, especially for the NRI (NRI = 0.3, SE 0.1, p = 0.007; IDI = 0.017, Table 4).

The prognostic accuracy of MR-proADM and the additional value it provides to the NEWS score was confirmed for *Deterioration Event*s and *Length of Stay* (Tables 4 and 5, and Figure 2 panels B and C).

Prognostic accuracy of MR-proADM in emergency departments

**Table 5.** Model comparisons. Outcomes 1, 2, and 3 refer to *Acuity Increase*, *Deterioration Event*, and *Length of Stay* respectively. The predictors are: *Set a* NEWS score alone; *Set b* NEWS score and MR-proADM; *Set c* NEWS score, MR-proADM, and other significant predictors and interactions detailed in Table 3.

| | AIC | Deviance | AUC (CI) or $R^2$ for linear regression | LR, (df) p-value | NRI (se), p-value | IDI (se), p-value |
|---|---|---|---|---|---|---|
| *Acuity Increase*: logistic regressions | | | | | | |
| Outcome 1 - predictor set *a.* | 348 | 356 | 0.55 (0.48, 0.62) | | | |
| Outcome 1 - predictor set *b.* | 343 | 353 | 0.61 (0.54, 0.69) | 5 (1), 0.033 | 0.3 (0.1), 0.007 | 0.017 (0.009), 0.058 |
| Outcome 1 - predictor set *c.* | 317 | 331 | 0.69 (0.63, 0.76) | 14 (2), 0.001* | 0.4 (0.1), 0.0004* | 0.05 (0.01),0.0009* |
| *Deterioration Event*: logistic regressions | | | | | | |
| Outcome 2 - predictor set *a.* | 199 | 207 | 0.57 (0.47, 0.68) | | | |
| Outcome 2 - predictor set *b.* | 192 | 202 | 0.65 (0.54, 0.76) | 7 (1), 0.007 | 0.4 (0.2), 0.003 | 0.04 (0.02), 0.10 |
| Outcome 2 - predictor set *c.* | 177 | 193 | 0.73 (0.63, 0.84) | 15 (3), 0.0019* | 0.5 (0.2), 0.012* | 0.06 (0.02), 0.0004* |
| *Length of Stay*: linear regressions (LR) | | | | | | |
| Outcome 3 - predictor set *a.* | 77 | -381 | 0.03 | | | |
| Outcome 3 - predictor set *b.* | 68 | -417 | 0.14 | 9 (1), <0.001 | | |
| Outcome 3 - predictor set *c.* | 67 | -420 | 0.16 | 1 (1), 0.026 | | |

**Note**: AIC = Akaike information criterion; AUC = area under the receiver operating characteristic curve; CI = 95% confidence interval; LR = likelihood ratio; df = degrees of freedom; NRI = net reclassification index; se = standard error; IDI = integrated discrimination improvement. * Comparison is between predictor set b. and c. Since there was a mismatch between the cases for predictor set a. and b. (10 missing values in COPD/HF), in the model with predictors set b. the 10 cases missing in predictor set c. were dropped to allow the comparison.

## Effect on prognostic accuracy when clinical information is added to the set of predictors

Secondary multivariable modelling evaluated the prognostic accuracy of MR-proADM when adjusted for the clinical factors in predictive set *c*: age, gender, CRP, WBC, presence of COPD or HF, presence of other comorbidities,

For *Acuity Increase,* COPD or HF comorbidity status and its interaction with MR-proADM level significantly improved the prognostic accuracy of the model: AUC increased from 0.61 (95% CI 0.54, 0.69) to 0.69 (95% CI 0.63, 0.76), likelihood ratio from 4 to 14, and net reclassification index from 0.3 to 0.4 (Table 5).

For *Deterioration Events*, the presence of other comorbidities (excluding COPD and HF) and $Age^2$ increased the prognostic accuracy of MR-proADM, (Table 4 and 5). The prognostic accuracy of *Length of Stay* (Outcome 3) of MR-proADM is also increased by including *Age* in the model (Table 4 and 5, Supplementary Figure s2).

## Potential confounding effects

**Shorter term outcomes**: NEWS and MR-proADM were less accurate in predicting *Acuity Increase* within 24 and 12 hours from admission than in predicting *Acuity Increase* within 72 hours (Supplementary Tables s1 and s2).

**Interval between admission NEWS scoring and blood collection:** Because ward processes did not allow the times of scoring NEWS and collecting blood to be specified for research, we assessed for a confounding effect from variation in the timings, but found no evidence for it (Supplementary Table s3).

**Correlations among biomarkers**. Diagnostic plots, shown in Supplementary Figures s2 and s3, show no multicollinearity in the data, no autocorrelation, no heteroscedasticity, and no data points that stood out in terms of their influence on results.

## Sensitivity and specificity

As overall measures of accuracy, sensitivity and specificity were calculated (where appropriate) for each model using Youden's index. The results are shown in

Supplementary Table s4. In practice, the trade-off between sensitivity and specificity would depend on the type of clinical decision to be made on the result (i.e. "rule-in" or "rule out") and this would differ from the approach in Youden's Index, which gives equal weight to false positive and false negative results.

### Internal Validation

C-statistic values after correcting for optimistic predictions were: for *Acuity Increase:* predictor set *a*, C-stat=0.53*;* predictor set b, C-stat=0.59; predictor set c, C-stat=0.66. For *Deterioration Events:* predictor set a, C-stat=0.52; predictor set b, C-stat=0.61, predictor set c, C-stat=0.68. For *Length of Stay:* predictor set a, $R^2$=0.003; predictor set b, $R^2$=0.12; predictor set c, $R^2$=0.13.

## Discussion

### Accuracy of prediction of deterioration by MR-proADM

This study shows that MR-proADM may be a clinically useful biomarker for predicting deterioration (i.e. *Acuity Increase*) within 72 hours from admission to hospital in patients with an admission NEWS score of 2 to 5. This contrasts with the performance of the NEWS score, assessed on admission, which did not predict deterioration within 72 hours, as might have been expected from previous evaluations [3-5 34]. This discrepancy with previous studies might be explained by differences in selection criteria for patients. Previous research included all patients admitted to ED, but our study selected patients with NEWS between 2 and 5, because a tool to predict deterioration would be most useful in this group.

For most of the observed *Acuity Increase* events, the reason for classification was an increase in the NEWS score. Because an increase in NEWS score reflects both measurement variation and physiological variation, additional exploratory analyses were carried out to assess the performance of MR-proADM with an operational definition of deterioration, *Deterioration Event*, designed to minimize measurement variation. NEWS on its own had low prognostic accuracy for *Deterioration Event*s. However, MR-proADM level, and NEWS score together predicted *Deterioration Events* with an AUC of 0.65. Considering baseline patient characteristics further increased the accuracy of the model (AUC = 0.73).

## Comorbidities and interactions with MR-proADM levels

MR-proADM levels in people with COPD and/or heart failure are chronically raised and are not predictive of deterioration. However, in other people whose MR-proADM levels are not chronically raised, high levels are predictive of *Acuity Increase* (Supplementary Figure s1). Including these comorbidities and their interaction with MR-proADM level increased the prognostic accuracy of the logistic regression model.

## Limitations

This study included only patients who were admitted with a NEWS score between 2 and 5. The prognostic accuracy of the MR-proADM would perhaps have been greater if more extreme cases had been included. However, patients with NEWS scores more than 5 are already known to be severely ill and to require close monitoring and/or management at higher levels of care.

Internal validation found that the uncorrected C-statistics are optimistic, which implies that external validation in an independent study would be useful. However, after correction for optimistic predictions, the study's conclusions remain unchanged.

## Interpretations and implications

The significance of MR-proADM in the prognostic models implies that it could provide additional prognostic information over and above NEWS score.

Secondary analyses suggest that a potentially useful clinical decision aid could be based on NEWS score, MR-proADM level, and clinical features.

## Future research and development

As a growing number of NHS hospitals are implementing the NEWS score on their clinical information systems, it should be practical to develop a decision aid based on admission NEWS score, MR-proADM level, and clinical features. Other biomarkers may further improve prognostic accuracy for deterioration, for example: lactate [3]; peroxiredoxin-4 (Prx4) and copeptin[22 35 36]; and soluble urokinase plasminogen activator receptor (suPAR)[37]. The feasibility, cost-effectiveness, and acceptability of such decision aids needs to be evaluated in further research.

A rapid point of care test for MR-proADM could facilitate the assessment process and reduce delays in arranging optimal levels of care and intensity of monitoring.

## Footnotes

**Contributors:** AJS, DAP, MP, and DS devised the study; RO, SG, and AJA managed the project; SG performed the statistical analyses with advice from DS; all authors contributed to the final manuscript.

**Acknowledgements**: The authors would like to thank the patients who accepted to be part of the study, and the team of research nurses for the data collection: Laura Shewan, Louise Taylor, Janine Gradwell, Karen Martin, Katherine Cullen, Gerry Jones, Graham Soulsby, Carmen Scott, and Helen Reed. We also thank the reviewers for their helpful suggestions and comments.

The NIHR Newcastle In Vitro Diagnostics Co-operative is funded by the NIHR.

**Funding:** The study was funded by B·R·A·H·M·S GmbH

**Competing interests**: None declared.

**Ethics and governance approvals:** The study was approved by the Newcastle & North Tyneside Research Ethics Committee (15/NE/0120), and by the R&D Committee of the Newcastle upon Tyne Hospitals NHS foundation Trust (reference number 7495).

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data are available

References

1. Royal College of Physicians of London. National early warning score (NEWS): standardising the assessment of acute-illness severity in the NHS—report of a working party. . London, 2012.

2. McGinley A, Pearse RM. A national early warning score for acutely ill patients. BMJ : British Medical Journal 2012;**345**.

3. Jo S, Yoon J, Lee JB, et al. Predictive value of the National Early Warning Score–Lactate for mortality and the need for critical care among general emergency department patients. Journal of Critical Care 2016;**36**:60-68.

4. Smith GB, Prytherch DR, Meredith P, et al. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation 2013;**84**(4):465-70.

5. Smith ME, Chiovaro JC, O'Neil M, et al. Early warning system scores for clinical deterioration in hospitalized patients: a systematic review. Annals of the American Thoracic Society 2014;**11**(9):1454-65.

6. Teasdale GM. National early warning score (NEWS) is not suitable for all patients. BMJ : British Medical Journal 2012;**345**.

7. Eccles SR, Subbe C, Hancock D, et al. CREWS: Improving specificity whilst maintaining sensitivity of the National Early Warning Score in patients with chronic hypoxaemia. Resuscitation 2014;**85**(1):109-11.

8. Badriyah T, Briggs JS, Meredith P, et al. Decision-tree early warning score (DTEWS) validates the design of the National Early Warning Score (NEWS). Resuscitation 2014;**85**(3):418-23.

9. Abbott TE, Vaid N, Ip D, et al. A single-centre observational cohort study of admission National Early Warning Score (NEWS). Resuscitation 2015;**92**:89-93.

10. Corfield AR, Lees F, Zealley I, et al. Utility of a single early warning score in patients with sepsis in the emergency department. Emergency medicine journal : EMJ 2014;**31**(6):482-7.

11. Bilben B, Grandal L, Søvik S. National Early Warning Score (NEWS) as an emergency department predictor of disease severity and 90-day survival in the acutely dyspneic patient – a prospective observational study. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine 2016;**24**:80.

12. Valenzuela-Sanchez F, Valenzuela-Mendez B, Rodriguez-Gutierrez JF, et al. New role of biomarkers: mid-regional pro-adrenomedullin, the biomarker of organ failure. Annals of translational medicine 2016;**4**(17):329.

13. Stokes NR, Dietz BW, Liang JJ. Cardiopulmonary laboratory biomarkers in the evaluation of acute dyspnea. Open access emergency medicine : OAEM 2016;**8**:35-45.

14. Sinning C, Zengin E, Zeller T, et al. Candidate biomarkers in heart failure with reduced and preserved ejection fraction. Biomarkers : biochemical indicators of exposure, response, and susceptibility to chemicals 2015;**20**(4):258-65.

15. Kutz A, Hausfater P, Amin D, et al. The TRIAGE-ProADM Score for an Early Risk Stratification of Medical Patients in the Emergency Department - Development Based on a Multi-National, Prospective, Observational Study. PloS one 2016;**11**(12).

16. Albrich WC, Dusemund F, Ruegger K, et al. Enhancement of CURB65 score with proadrenomedullin (CURB65-A) for outcome prediction in lower respiratory tract infections: derivation of a clinical algorithm. BMC infectious diseases 2011;**11**:112.

17. Lopez J, Martinez A. Cell and molecular biology of the multifunctional peptide, adrenomedullin. International review of cytology 2002;**221**:1-92.

18. Serrano-Ponz M, Rodrigo-Gasque C, Siles E, et al. Temporal profiles of blood pressure, circulating nitric oxide, and adrenomedullin as predictors of clinical outcome in acute ischemic stroke patients. Molecular medicine reports 2016;**13**(5):3724-34.

19. Zudaire E, Portal-Nunez S, Cuttitta F. The central role of adrenomedullin in host defense. Journal of leukocyte biology 2006;**80**(2):237-44.

20. Zuur-Telgen MC, Brusse-Keizer MG, Vandervalk PD, et al. Stable state MR-proadrenomedullin level is a strong predictor for mortality in COPD patients. Chest 2013.

21. Pedowska-Wloszek J, Kostrubiec M, Kurnicka K, et al. Midregional proadrenomedullin (MR-proADM) in the risk stratification of patients with acute pulmonary embolism. Thrombosis research 2013;**132**(5):506-10.

22. Nickel CH, Messmer AS, Geigy N, et al. Stress markers predict mortality in patients with nonspecific complaints presenting to the emergency department and may be a useful risk stratification tool to support disposition planning. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2013;**20**(7):670-9.

23. Landman GW, van Dijk PR, Drion I, et al. Midregional fragment of proadrenomedullin, new-onset albuminuria, and cardiovascular and all-cause mortality in patients with type 2 diabetes (ZODIAC-30). Diabetes care 2014;**37**(3):839-45.

24. Funke-Kaiser A, Mann K, Colquhoun D, et al. Midregional proadrenomedullin and its change predicts recurrent major coronary events and heart failure in stable coronary heart disease patients: the LIPID study. International journal of cardiology 2014;**172**(2):411-8.

25. Gaggin HK, Januzzi JL, Jr. Biomarkers and diagnostics in heart failure. Biochimica et biophysica acta 2013;**1832**(12):2442-50.

26. Vigue B, Leblanc PE, Moati F, et al. Mid-regional pro-adrenomedullin (MR-proADM), a marker of positive fluid balance in critically ill patients: results of the ENVOL study. Critical care (London, England) 2016;**20**(1):363.

27. Tzikas S, Keller T, Ojeda FM, et al. MR-proANP and MR-proADM for risk stratification of patients with acute chest pain. Heart (British Cardiac Society) 2013;**99**(6):388-95.

28. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. Journal of clinical epidemiology 1996;**49**(12):1373-9.

29. van Smeden M, de Groot JAH, Moons KGM, et al. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. BMC Medical Research Methodology 2016;**16**(1):163.

30. Travaglino F, Russo V, De Berardinis B, et al. Thirty and ninety days mortality predictive value of admission and in-hospital procalcitonin and mid-regional pro-adrenomedullin testing in patients with dyspnea. Results from the VERyfing DYspnea trial. The American journal of emergency medicine 2014;**32**(4):334-41.

31. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria.: R Foundation for Statistical Computing, 2011.

32. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., et al. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Statistics in medicine 2008;**27**(2):157-72; discussion 207-12.

33. Steyerberg EW, Vedder MM, Leening MJ, et al. Graphical assessment of incremental value of novel markers in prediction models: From statistical to decision analytical perspectives. Biometrical journal Biometrische Zeitschrift 2015;**57**(4):556-70.

34. Alam N, Vegting IL, Houben E, et al. Exploring the performance of the National Early Warning Score (NEWS) in a European emergency department. Resuscitation 2015;**90**:111-5.

35. Nickel CH, Ruedinger J, Misch F, et al. Copeptin and peroxiredoxin-4 independently predict mortality in patients with nonspecific complaints presenting to the emergency department. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2011;**18**(8):851-9.

36. Iversen K, Gotze JP, Dalsgaard M, et al. Risk stratification in emergency patients by copeptin. BMC medicine 2014;**12**:80.

37. Rasmussen LJ, Ladelund S, Haupt TH, et al. Soluble urokinase plasminogen activator receptor (suPAR) in acute care: a strong marker of disease presence and severity, readmission and mortality. A retrospective cohort study. Emergency medicine journal : EMJ 2016.

## Figure legends

**Figure 1.** Patient recruitment process.

**Figure 2. Panel A.** Prognostic accuracy for *Acuity Increase*; predictor set *a*: NEWS; predictors set *b*: NEWS, MR-proADM; predictor set c: NEWS, MR-proADM, COPD/HF, interaction between MR-proADM and COPD/HF. **Panel B.** Comparisons as for panel **A** but for predicting a *Deterioration Event*; predictor set *c:* NEWS score, MR-proADM level, $Age^2$, other comorbidities. **Panel C.** *Length of Stay* predicted by MR-proADM level.

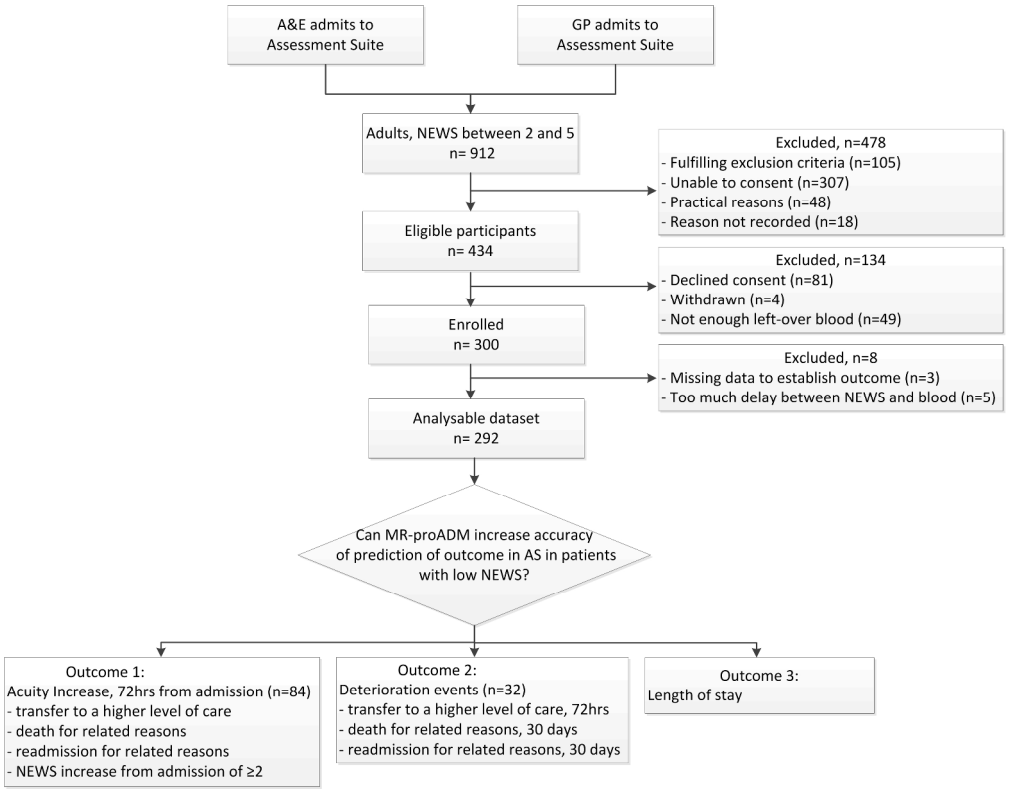



Figure 1. Patient recruitment process.

255x200mm (300 x 300 DPI)

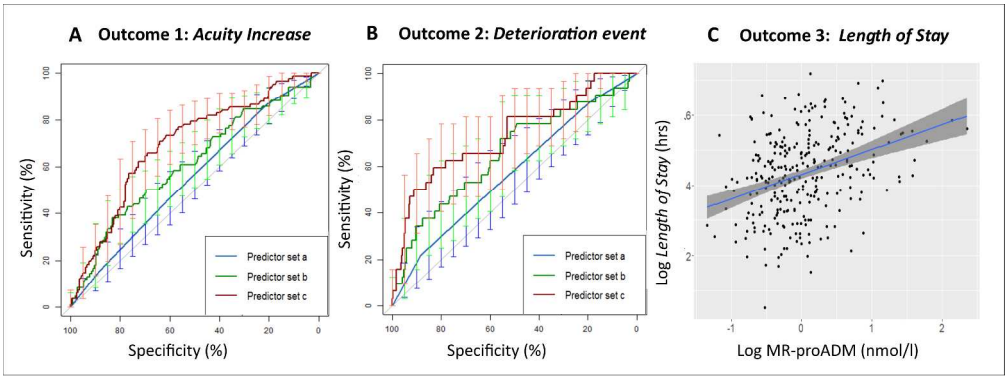



Figure 2. Panel A. Prognostic accuracy for Acuity Increase; predictor set a: NEWS; predictors set b: NEWS, MR-proADM; predictor set c: NEWS, MR-proADM, COPD/HF, interaction between MR-proADM and COPD/HF. Panel B. Comparisons as for panel A but for predicting a Deterioration Event; predictor set c: NEWS score, MR-proADM level, Age2, other comorbidities. Panel C. Length of Stay predicted by MR-proADM level.

Link text : Figure 2

320x118mm (300 x 300 DPI)

Supplementary information for Graziadio et al, 2018

Supplementary information for *Can mid-regional pro-adrenomedullin (MR-proADM) increase the prognostic accuracy of NEWS in predicting deterioration in patients admitted to hospital with mild to moderately severe illness? A prospective single-centre observational study*. Graziadio et al, 2018

## Table of Contents

Supplementary information for Graziadio et al, 2018

# Additional information on Methods

## NEWS as a predictor of deterioration in patients with mild to moderately severe illness

### Operational definitions of deterioration

In the original validation of NEWS, about 2% of the population had the combined outcome of cardiac arrest, unanticipated ICU admission, or death — each within 24 hours [Smith 2013]. Furthermore, the proportions for each of the three individual outcomes and the composite outcomes increased monotonically through the range of NEWS scores.

Thus, as designed in its development, a NEWS score between 2 and 5 defines a population at low risk of cardiac arrest within 24 hours, death within 24 hours, or ICU admission within 24 hours.

However, a clinically important proportion of patients admitted to A&E or Medical Admissions Unit with mild to moderately severe illness (NEWS scores between 2 and 5) do deteriorate, and the NEWS score is, by design, not able to identify these patients.

The improvement challenge is to identify biomarkers that will increase the discrimination of low NEWS scores. And the methodological challenge was to develop convenient and effective operational definitions of deterioration from mild/moderately severe.

As NEWS is used to monitor changes in severity of illness, we decided to base our primary operational definition of deterioration on an increase of at least 2 in the NEWS score.

***Acuity Increase***. The primary outcome was the proportion of patients who, within 72 hours, had any combination of:

- an increase of at least 2 in the NEWS score
- transfer to a higher-dependency bed or monitored area
- death
- for those discharged from hospital, re-admission for medical reasons.

We labelled this measure *Acuity Increase*.

Because there was concern about variations in NEWS scoring and about using NEWS to predict a change in NEWS (which it is designed not to do in this study's population), we defined two other measures of deterioration, one direct, and the other indirect***:***

***Deterioration Event***: the occurrence of one or more of the following:

- transfer to higher level of care within 72 hours from admission;
- death (for reasons related to admission) within 30 days;
- re-admission to hospital (for the same reason as the previous admission) within 30 days from first admission

***Length of Stay***: the duration in days from admission to discharge or death

Supplementary information for Graziadio et al, 2018

### Predictors of deterioration for statistical modelling

In the analyses, we included NEWS as a possible predictor of deterioration. As expected, NEWS scores consistently do not predict deterioration, for all three of our operational definitions.

In line with the original validation of NEWS, we included NEWS as an ordinal variable [Smith 2013].

## Analytical data exploration

Univariate logistic regressions were used to investigate whether the relationship between outcome variables (i.e. deterioration measures) and the input variables (NEWS and MR-proADM, age, comorbidities, gender, CRP, and WBC) were linear. If they were not linear, log transformation and squared transformation were applied. If the transformation substantially lowered the AIC, then the transformed variable was used in statistical analyses.

For categorical variables with multiple ordinal levels (i.e. NEWS score), the univariate analysis informed if it was appropriate to include the variable in the model as a continuous or categorical factor. If the coefficients in the univariate models increased linearly, then a linear relationship with the outcome could be assumed, and the variable was included in the model as continuous, otherwise the variable was treated as a categorical factor.

Univariate analyses were also used to identify the variables that affected the outcome significantly. Variables with a probable relationship with the outcome variable ($p<0.1$) were included in the full model logistic regression.

## Visual data exploration and interaction between MR-proADM and COPD/HF

After formatting the datasets, all variables were graphed (bar-charts for categorical variables, and scatterplots/histograms for continuous variables) and visually checked for outliers and distributions that seemed potentially erroneous.

If outliers were identified, the cause(s) were investigated to understand whether they were due to human error or they were genuine data. Outliers were kept in the primary analysis. In a secondary sensitivity analysis, outliers were removed and the same analyses repeated to assess the impact on the results. If the coefficients of the predictors changed substantially, both models would be described. There was one genuine outliner patient with a very high level of MR-proADM compared to the population mean, but its exclusion made no meaningful difference to the results, and the subject was included in the final analysis.

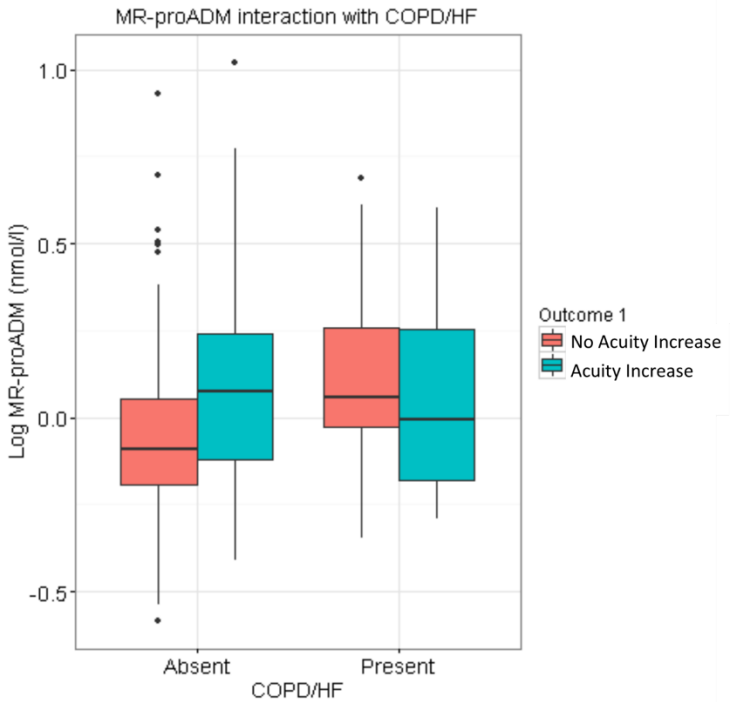
The influence of potentially important factors (such as comorbidities) on the ability of the MR-proADM to predict deterioration was explored graphically.

### Interaction between MR-proADM and COPD/HF

A significant interaction between MR-proADM and the presence of COPD/HF was discovered, and therefore included in the logistic regression (Outcome 1, predictor set c). The plot is

shown in **Supplementary Figure s1**. This interaction showed that the MR-proADM level was increased in patients who deteriorated, but only if they did not have COPD or HF.

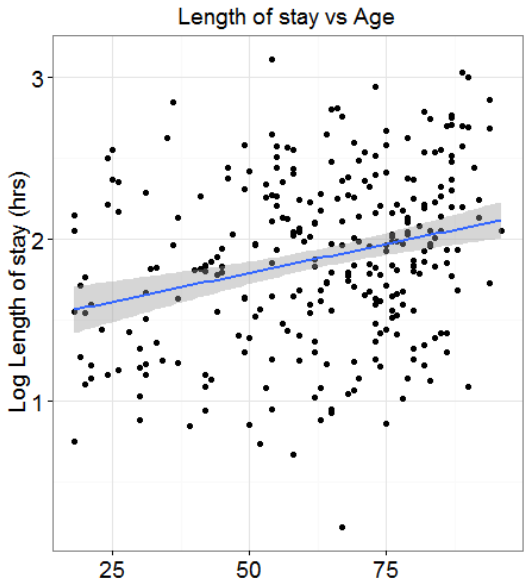
There was no suggestion that age; comorbidities: COPH and HF; other comorbidities; CRP; or WBC would improve the accuracy of prediction.




Interaction between MR-proADM and COPD/HF.

## Relationship between Length of Stay and Age

In **Supplementary Figure s1** the relationship between *Length of Stay* and *Age* is shown.

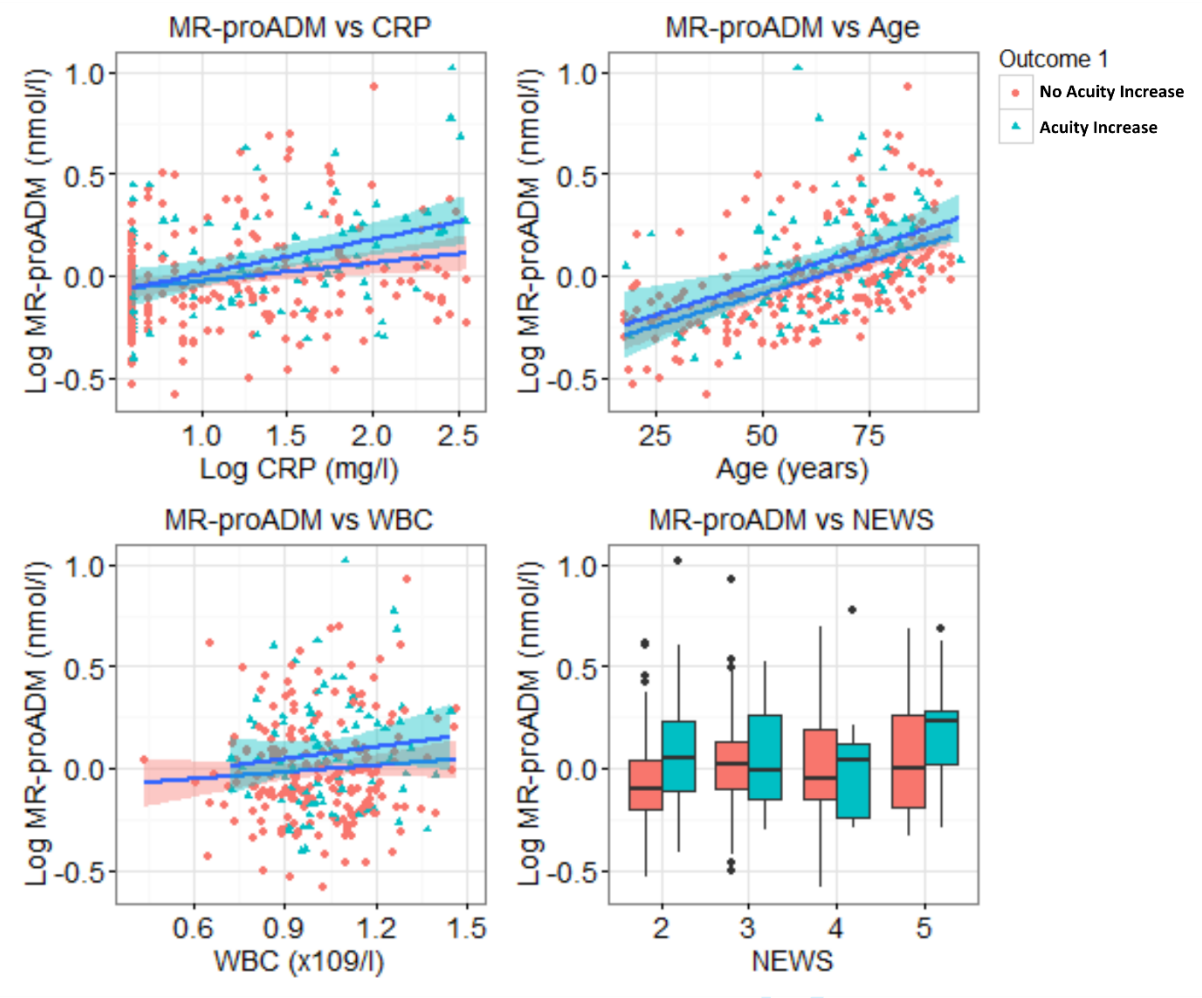



**Supplementary Figure s1**. Relationship between *Length of Stay* and *Age*.

Supplementary information for Graziadio et al, 2018

## Checking for multicollinearity and autocorrelation

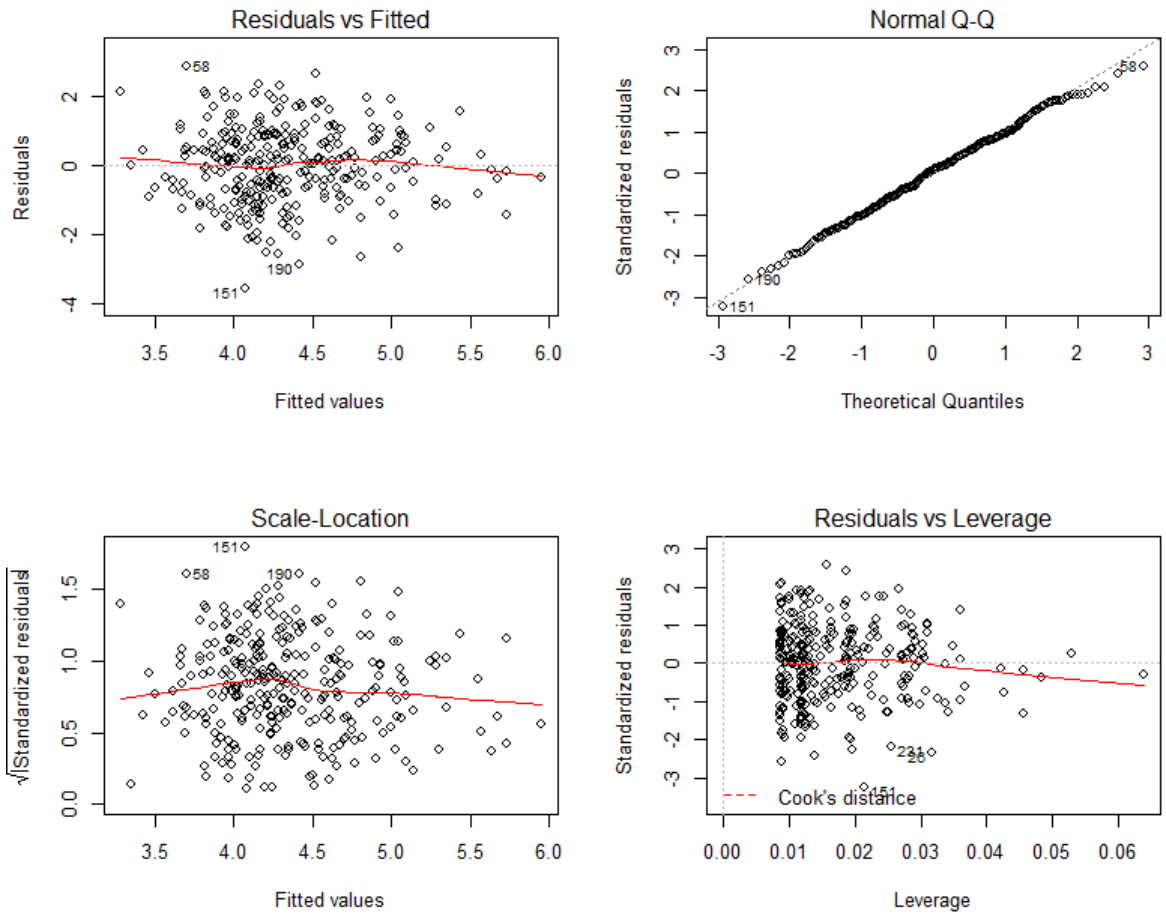
Correlations among biomarkers were also investigated through plotting to evaluate multicollinearity and added value of MR-proADM versus other biomarkers. Plots are shown in **Supplementary Figure s2**.




**Supplementary Figure s2.** Associations between MR-proADM and CRP, age, WBC, and NEWS.

Supplementary information for Graziadio et al, 2018

## Diagnostic plots for length of stay analysis




**Supplementary Figure s3.** Diagnostic plots for linear regression evaluating the prediction accuracy of MR-proADM for *Length of Stay*.

The diagnostics of the model showed no multicollinearity in the data since all the correlation coefficients among the independent variables were smaller than 0.5. No autocorrelation was found in the data, thus residuals are independent from each other: the Durbin-Watson test estimated d = 2.02 (p = 0.56). Evidence for homoscedasticity was provided graphically by the randomly scattered points and almost horizontal fitted lines in **Supplementary Figure s3**, (Residuals vs fitted plot). Analysis of Cook's distance showed that there were no influential points (d <4/51, **Supplementary Figure s3**).

## Analyses of shorter term outcomes

The analyses found that NEWS and MR-proADM had much lower accuracy in predicting *Acuity Increase* at 24 and 12 hours from admission than in predicting *Acuity Increase* at 72 hours as apparent in **Supplementary Table s1**.

Supplementary information for Graziadio et al, 2018

**Supplementary Table s1**. Logistic regression for NEWS and MR-proADM predicting *Acuity Increase* within 24 hours. AIC = 326; AUC = 0.59.

| Covariate | Beta | CI | Odds ratio(CI) | P-value |
|---|---|---|---|---|
| **Intercept** | –1.42 | –1.70, –0.64 | 0.42 (0.27, 0.64) | NA |
| **NEWS 3** | –0.006 | –0.64, 0.71 | 1.04 (0.53, 2.04) | 0.985 |
| **NEWS 4** | –0.37 | –1.58, 0.20 | 0.52 (0.21, 1.22) | 0.368 |
| **NEWS 5** | 0.21 | –0.66, 1.01 | 1.25 (0.52, 2.91) | 0.6244 |
| **MR-proADM** | 0.22 | –0.06, 0.44 | 1.20 (0.95, 1.55) | 0.0547 |

**Supplementary Table s2.** Logistic regression for NEWS and MR-proADM predicting *Acuity Increase* within 12 hours. AIC = 266; AUC = 0.57.

| Covariate | Beta | CI | Odds ratio(CI) | P-value |
|---|---|---|---|---|
| **Intercept** | –1.83 | –2.47, –1.24 | 0.16 (0.08, 0.29) | NA |
| **NEWS 3** | 0.29 | –0.46, 1.04 | 1.34 (0.63, 2.85) | 0.442 |
| **NEWS 4** | –0.15 | –1.57, 0.76 | 0.86 (0.31, 2.14) | 0.756 |
| **NEWS 5** | 0.29 | –0.74, 1.23 | 1.33 (0.48, 3.42) | 0.564 |
| **MR-proADM** | 0.06 | –0.24, 0.31 | 1.06 (0.79, 1.36) | 0.656 |

## Analyses of time-lag effect between news assessment and blood collection for assessment of MR-proADM levels

Given the practicalities involved, it was not possible to stipulate the timings of taking the NEWS on admission and collecting the blood sample for MR-proADM testing. It was expected that difference in times would normally be less than 6 hours, but in 44 subjects the time difference was more than 6 hours.

To investigate the impact of time differences being greater than expected, another analysis was carried out excluding subjects for whom the difference was more than 6 hours (time-lag compliant dataset). The hypothesis was that, if the time difference was an important parameter for the predictive accuracy of MR-proADM level, model coefficients would be greater and confidence intervals narrower for the compliant model. This was not the case; results were similar in the full dataset with 292 subjects and in the compliant dataset with 248 subjects (**Supplementary Table s3**).

Supplementary information for Graziadio et al, 2018

**Supplementary Table s3.** Logistic regression for NEWS and MR-proADM predicting *Acuity Increase* for the time-lag compliant dataset. AIC = 295; AUC = 0.60.

| Covariate | Beta | CI | Odds ratio(CI) | P-value |
|---|---|---|---|---|
| **Intercept** | –1.15 | –1.70, –0.64 | 0.42 (0.27, 0.64) | NA |
| **NEWS 3** | –0.04 | –0.64, 0.71 | 1.04 (0.53, 2.04) | 0.909 |
| **NEWS 4** | –0.65 | –1.58, 0.20 | 0.52 (0.21, 1.22) | 0.152 |
| **NEWS 5** | 0.22 | -0.66, 1.01 | 1.25 (0.52, 2.91) | 0.613 |
| **MR-proADM** | 0.19 | -0.06, 0.44 | 1.20 (0.95, 1.55) | 0.135 |

## Estimation of sensitivities and specificities

For completeness we estimated the sensitivity and specificity for each model in the article (**Supplementary Table s4**). We used the Youden's index to estimate the cut-off. In the next phase of the MR-proADM evaluation, the cut-off will be re-estimated through a decision analysis informed by the role of the test in the pathway.

**Supplementary Table s4.** Sensitivity and specificity of the logistic regression models. Predictor set a. was excluded from the table since the AUROC was too low to calculate meaningful diagnostic accuracy data.

| Models | Sensitivity | Specificity |
|---|---|---|
| ***Acuity Increase*** | | |
| **Predictor set b.** | 0.38 (0.29, 0.49) | 0.83 (0.78, 0.88) |
| **Predictor set c.** | 0.66 (0.55, 0.76) | 0.69 (0.63, 0.76) |
| ***Deterioration Event*** | | |
| **Predictor set b.** | 0.72 (0.56, 0.88) | 0.56 (0.50, 0.62) |
| **Predictor set c.** | 0.59 (0.44, 0.75) | 0.83 (0.78, 0.87) |

TRIPOD Checklist: Prediction Model Development

| Section/Topic | Item | Checklist Item | Page |
|---|---|---|---|
| **Title and abstract** | | | |
| Title | 1 | Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted. | P1 L1-3 |
| Abstract | 2 | Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions. | P2 L22-46 |
| **Introduction** | | | |
| Background and objectives | 3a | Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models. | P4-5 L52-84 |
| | 3b | Specify the objectives, including whether the study describes the development or validation of the model or both. | P5 L82-84 |
| **Methods** | | | |
| Source of data | 4a | Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable. | P5 L87 |
| | 4b | Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up. | P5 L87-91 |
| Participants | 5a | Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres. | P5 L87-97 |
| | 5b | Describe eligibility criteria for participants. | P5 L98-102 |
| | 5c | Give details of treatments received, if relevant. | N/A |
| Outcome | 6a | Clearly define the outcome that is predicted by the prediction model, including how and when assessed. | P6-7 L128-149 |
| | 6b | Report any actions to blind assessment of the outcome to be predicted. | P6 L123-126 |
| Predictors | 7a | Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured. | P4 L54-57 P4-5 L70-84 P5-6 L106-117 |
| | 7b | Report any actions to blind assessment of predictors for the outcome and other predictors. | P6 L123-126 |
| Sample size | 8 | Explain how the study size was arrived at. | P5 L92-97 |
| Missing data | 9 | Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method. | P8-9 L197-201 |
| Statistical analysis methods | 10a | Describe how predictors were handled in the analyses. | P7-8 L151-194 |
| | 10b | Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation. | P7-8 L151-194 P7 L153-155 |
| | 10d | Specify all measures used to assess model performance and, if relevant, to compare multiple models. | P7-8 L156-194 |
| Risk groups | 11 | Provide details on how risk groups were created, if done. | N/A |
| **Results** | | | |
| Participants | 13a | Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful. | P10-11 Table 1. |
| | 13b | Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome. | P10-11 Table 1. |
| Model development | 14a | Specify the number of participants and outcome events in each analysis. | P10-11 Table 1. |
| | 14b | If done, report the unadjusted association between each candidate predictor and outcome. | P13-15 Table3&4 |
| Model specification | 15a | Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point). | P13-15 Table3&4 |
| | 15b | Explain how to the use the prediction model. | P15 L244-253 |
| Model performance | 16 | Report performance measures (with CIs) for the prediction model. | P13-15 Table3&4 P15 L244-253 |
| **Discussion** | | | |
| Limitations | 18 | Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data). | P18-19 L306-316 |
| Interpretation | 19b | Give an overall interpretation of the results, considering objectives, limitations, and results from similar studies, and other relevant evidence. | P17-19 L278-323 |
| Implications | 20 | Discuss the potential clinical use of the model and implications for future research. | P19 L325-335 |
| **Other information** | | | |
| Supplementary | 21 | Provide information about the availability of supplementary resources, such as study | See Comments |

TRIPOD Checklist: Prediction Model Development

| | | | |
|---|---|---|---|
| information | | protocol, Web calculator, and data sets. | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study. | P20 L346 |

Comments:

1. **Item 21: Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.**

Supplementary material - with additional information on methods and results - is attached as separate document. Study protocol and data sets will be available in due course, new project website currently under construction.

BMJ Open

# Can mid-regional pro-adrenomedullin (MR-proADM) increase the prognostic accuracy of NEWS in predicting deterioration in patients admitted to hospital with mild to moderately severe illness? A prospective single-centre observational study.



| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2017-020337.R2 |
| Article Type: | Research |
| Date Submitted by the Author: | 16-Aug-2018 |
| Complete List of Authors: | Graziadio, Sara; Newcastle University, Institute of Cellular Medicine; Newcastle Upon Tyne Hospitals NHS Foundation Trust, NIHR Newcastle In Vitro Diagnostics Co-operative<br>Price, David; Newcastle Upon Tyne Hospitals NHS Foundation Trust, Department of Infectious Diseases; Newcastle Upon Tyne Hospitals NHS Foundation Trust, NIHR Newcastle In Vitro Diagnostics Co-operative<br>O'Leary, Rachel; Newcastle Upon Tyne Hospitals NHS Foundation Trust, NIHR Newcastle In Vitro Diagnostics Co-operative; Newcastle University, Institute of Cellular Medicine<br>Stocken, Deborah; University of Leeds, Leeds Institute of Clinical Trials Research; Newcastle University, Institute of Health and Society<br>Power, Michael; Newcastle Upon Tyne Hospitals NHS Foundation Trust, NIHR Newcastle In Vitro Diagnostics Co-operative; Newcastle University , Institute of Cellular Medicine<br>Allen, A; Newcastle University, Institute of Cellular Medicine; Newcastle Upon Tyne Hospitals NHS Foundation Trust, NIHR Newcastle In Vitro Diagnostics Co-operative<br>Simpson, A John; Newcastle University, Institute of Cellular Medicine; Newcastle Upon Tyne Hospitals NHS Foundation Trust, NIHR Newcastle In Vitro Diagnostics Co-operative |
| Primary Subject Heading: | Diagnostics |
| Secondary Subject Heading: | Evidence based practice, Emergency medicine |
| Keywords: | INTERNAL MEDICINE, ACCIDENT & EMERGENCY MEDICINE, Adult intensive & critical care < INTENSIVE & CRITICAL CARE |
| | |

Prognostic accuracy of MR-proADM in emergency departments

**Title:** Can mid-regional pro-adrenomedullin (MR-proADM) increase the prognostic accuracy of NEWS in predicting deterioration in patients admitted to hospital with mild to moderately severe illness? A prospective single-centre observational study.

**Authors:** Sara Graziadio[1, 2], D. Ashley Price[1, 3], Rachel A. O'Leary [1, 2], Deborah D. Stocken[4, 5], Michael Power[1, 2], A. Joy Allen[1, 2], A. John Simpson[1, 2]

[1] NIHR Newcastle In Vitro Diagnostics Cooperative, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, NE1 4LP, UK

[2] Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK

[3] Department of Infectious Diseases, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, NE1 4LP, UK

[4] Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, LS2 9JT, UK

[5] Institute of Health and Society, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK

**Corresponding author:** Dr Sara Graziadio, NIHR Newcastle In Vitro Diagnostics Cooperative, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, NE1 4LP, UK. Sara.Graziadio@ncl.ac.uk.

Word counts

Abstract: 275

Paper (including tables, figures, legends, and references): 6150

Prognostic accuracy of MR-proADM in emergency departments


## Abstract

**Objective** To assess the value added to the NEWS score by MR-proADM blood level in predicting deterioration in mild to moderately ill people.

**Design** Prospective observational study

**Setting** The Medical Admissions Suite of the Royal Victoria Infirmary, Newcastle.

**Participants** 300 adults with NEWS score between 2 and 5 on admission. Exclusion criteria included receiving palliative care, or admitted for social reasons or self-harming. Patients were enrolled between September and December 2015, and followed-up for 30 days after discharge.

**Outcome measure** The primary outcome measure was the proportion of patients who, within 72 hours, had an *Acuity Increase*, defined as any combination of: an increase of at least 2 in the NEWS score; transfer to a higher-dependency bed or monitored area; death; or for those discharged from hospital, re-admission for medical reasons.

**Results** NEWS and MR-proADM together predicted *Acuity Increase* more accurately than NEWS alone, increasing the AUC to 0.61 (95% CI 0.54, 0.69) from 0.55 (95% CI 0.48, 0.62). When the confounding effects of presence of chronic obstructive pulmonary disease or heart failure and interaction with MR-proADM were included, the prognostic accuracy further increased the AUC to 0.69 (95% CI 0.63, 0.76).

**Conclusions** MR-proADM is potentially a clinically useful biomarker for deterioration in patients admitted to hospital with a mild to moderately severe acute illness, i.e. with NEWS score between 2 and 5. As a growing number of NHS hospitals are routinely recording the NEWS score on their clinical information systems, further research should assess the practicality and utility of developing a decision aid based on admission NEWS score, MR-proADM level, and possibly other clinical data and other biomarkers that could further improve prognostic accuracy.



## Keywords

Biochemistry, diagnosis, health services research

## Strengths and limitations of this study

- This is the first study to use rigorous statistical methods to assess the value added by MR-proADM to the admission NEWS score for predicting clinically important deterioration in mild to moderately ill patients.
- Overall prognostic accuracy might have been greater had more severely ill patients been included, but the aim of this study was to predict deterioration in less severely ill patients who could benefit from closer observation.
- This was an observational study, and thus could not directly assess the utility of more accurate prediction of deterioration.
- Initial evidence for MR-proADM as a biomarker for deterioration appears promising, but requires further validation for clinical utility.

## Introduction

The National Early Warning Score (NEWS) is recommended for assessing severity of illness in patients presenting in primary or secondary NHS care and for surveillance of patients in hospital [1 2]. Six physiological parameters (which can be measured at the bedside) are scored: respiratory rate, oxygen saturation, temperature, systolic blood pressure, pulse rate, and level of consciousness. The scores are aggregated, and, if the patient requires oxygen, the total is increased. NEWS predicts death, cardiac arrest, and unplanned intensive care unit (ICU) admission within 24 hours [3-5]. However, NEWS does not identify all patients who turn out to be seriously ill [6-8], and there are also patients whose NEWS score is usually elevated and who do not require the level of observation that the NEWS tool would suggest. For example, people with chronic obstructive pulmonary disease (COPD) or chronic heart failure (HF) have higher baseline NEWS scores than those without these comorbidities. The prognostic accuracy of NEWS for patients presenting to the Emergency Department (ED) has been confirmed in a wide range of severity of illness [9 10], as has its reduced accuracy in people with COPD [11]. But, no previous studies of the prognostic accuracy of NEWS in the ED/Medical Admissions Unit (MAU) have focussed on patients admitted with mild to moderately severe illness. Since a clinically important proportion of these patients do deteriorate unexpectedly, improved risk stratification would be useful.

Mid-regional pro-adrenomedullin (MR-proADM) is one of several promising biomarkers for severe illness and deterioration [12-16].

MR-proADM is a precursor of adrenomedullin (ADM), a member of the calcitonin peptide family. ADM is widely expressed and has roles in vasodilation, immune modulation, and metabolic regulation. It is up-regulated in severe infections, inflammation, vasodilation, stimulation of diuresis, increased cardiac output, and stroke [17-19]. ADM has a short half-life, but MR-proADM is more stable and directly reflects ADM concentrations in blood. Both ADM and MR-proADM levels are strongly associated with risk of mortality, regardless of aetiology [20-26] . In people presenting with acute chest pain, MR-proADM has been reported to improve the Global Registry of Acute Coronary Events risk classification by 41% [27]. As with the NEWS score, people with COPD or chronic heart failure have higher baseline levels of MR-proADM.

The aim of this study was to assess whether the MR-proADM level used alongside the NEWS score would improve prediction of deterioration over NEWS score alone in patients admitted to the MAU with mild to moderately severe illness.

# Methods

## Patient and Public Involvement

Patients and the public were not specifically involved in the planning and execution of this study. However, the NIHR now requires that the research it supports includes active involvement and engagement with patients and the public.

## Study participants and study design

This was a prospective observational cohort study. Patients were enrolled between September and December 2015 at the Royal Victoria Infirmary, Newcastle, and followed-up for 30 days after discharge. If the patient died within the 30 days of follow up, this and the cause of death were recorded. Adults admitted to the MAU were recruited for the study between 9 am and 4 pm on weekdays.

Sample size was based on a pragmatic recruitment target for a three-month observational study. A recent unpublished audit conducted in the MAU at the Royal Victoria Infirmary found a deterioration rate of 20%. With 300 patients and complete data collection, 60 events would be anticipated. With this number of events, a multivariable prediction model could include up to six independent predictors. This is based on a widely accepted rule of thumb that models with fewer than ten events per predictor tend to be over-fitted [28]. However, recent research suggests that the "ten events per variable" rule of thumb may be optimistic [29]. Because the aim of this study was to assess if further research would be indicated, 60 is considered an acceptable number of events, even if the rule of thumb is optimistic.

Patients were considered eligible for inclusion in the study if their NEWS score on admission was at least 2 and not greater than 5, and all NEWS parameters were recorded. Patients were excluded from the study if they were receiving palliative care, were admitted for social reasons only, or were self-harming, or overdosing with drugs or other substances.

All participants provided written informed consent, and the study was approved by the Newcastle & North Tyneside Research Ethics Committee (15/NE/0120).

## Recorded data

Demographic and admission data included: gender, year of birth, reason for admissions, diagnosis on discharge, and the presence of comorbidities in which baseline MR-proADM levels are chronically raised: COPD with hypoxia ($PaO_2$ <10 kPa) [7]; HF [30]; acute brain injury [6]; acute coronary syndrome [27]; acute venous thromboembolism [21]; high International Normalized Ratio (INR>2); acute kidney injury; electrolyte disturbances ($Na^+$ <130 or >150 mmol/L; $K^+$ <3.0 or >5.5 mmol/L); hyperglycaemia in type 1 diabetes (random glucose >10 mmol/L).

The NEWS score was assessed at on admission and over the next 72 hours, and the scores and assessment times recorded. The 7 clinical parameters used to determine the NEWS score were recorded for the baseline (admission) assessment only. Baseline NEWS scores were used to determine eligibility for this study. Subsequent NEWS scores were used in the analyses to identify deterioration.

Blood samples were taken at hospital admission for assessment of MR-proADM, C-Reactive Protein (CRP) and white blood count (WBC).

## Laboratory tests

Plasma was obtained from blood samples (collected in ethylenediaminetetraacetic acid, EDTA) that were no longer clinically required. Plasma was stored in aliquots at –80º C.

MR-proADM was assayed in the on-site Blood Sciences Laboratory using the B R A H M S Kryptor system according to the manufacturer's instructions.

Blood samples were analysed in batches by personnel blinded with regard to the condition and NEWS score of the patient. Nurses who assessed the NEWS score and healthcare professionals managing patients in the MAU were blinded to MR-proADM results.

## Outcome measures

**Outcome measure 1: *Acuity Increase***. A patient was classified as having an *Acuity Increase* if one or more of the following occurred within 72 hours from admission:

1. transfer to a higher level of care (ICU or high dependency unit)
2. readmission to hospital for reasons related to the initial admission
3. death for reasons related to the initial admission
4. NEWS score increased by at least two compared to the admission score

**Outcome measure 2: *Deterioration Event***. For most of the observed *Acuity Increase* cases the reason for classification was an increase in the NEWS score (Table 1). An increase in NEWS score reflects both measurement variation and physiological variation, so additional exploratory analyses were carried out to assess the performance of MR-proADM in predicting deterioration. *Deterioration Events* were classified as the occurrence of one or more of the following:

1. transfer to higher level of care within 72 hours from admission;
2. death (for reasons related to the admission) within 30 days;
3. re-admission to hospital (for the same reason as the previous admission) within 30 days from first admission.

Classification based on this definition is unlikely to be subject to clinically important measurement variation. This analysis, therefore, should optimise the prognostic accuracy for events which are both clinically and economically important.

**Outcome measure 3: *Length of Stay***. *Length of Stay* was defined as the duration (in days) from admission to discharge or death.

## Statistical analysis

All data analyses were performed using the R language, version 3.2.0 [31], with the support of RStudio, version 0.99.896 (RStudio, Inc). The following R packages were used: ggplot2, pROC, psych, PredictABEL, Hmisc, rms.

Logistic regression models were compared for their accuracy in predicting deterioration outcome measures as pre-specified in an analysis plan. Analyses are presented as unadjusted parameter estimates of risk (odds ratio (OR), with confidence intervals (CI)) and estimates adjusted for identified clinical confounding factors. The aims of the multivariable analyses were twofold: first, to estimate the effect size and significance adjusted for other identified influential predictors and interactions; second, to investigate whether the addition of other predictors improved the goodness of fit and accuracy of prediction.

Only complete cases were analysed since missingness was minimal: 10 records without data on co-morbidities (details in footnote in Table 1).

For each measure of deterioration (*Acuity Increase, Deterioration Event*, and *Length of Stay*), logistic regression models were compared for the following sets of predictor variables:

*Predictor set a.* Comparator (base case): NEWS score on admission

*Predictor set b.* Primary analysis: NEWS score, MR-proADM

*Predictor set c.* Secondary analyses: NEWS score and MR-proADM always included. Age, gender, CRP, WBC, presence of COPD or HF, presence of other comorbidities, and interactions between predictors when appropriate.

Predictors (and the underlying assumption of linearity of their relationship with the outcome of interest) were initially investigated through univariate analyses based on simple log and quadratic functions. Transformations were applied if they improved the goodness of fit as assessed by the Akaike information criterion (AIC), and were retained in the multivariable setting. NEWS was treated as an ordinal variable. We assessed interactions through visual data exploration without significance testing as the study was not powered for this. For the multivariable regression models, the set of predictors was assessed for independence through backward elimination, based on changes in AIC.

The analysis plan for the secondary outcome of *Length of Stay* was similar: using multiple linear regressions based on transformed outcomes to address non-normality. Dependent and exploratory variables were log-transformed if not normally distributed. Normality was assessed by visualizing the data. More details on the methods used are reported in the online Supplemental Material.

Goodness of fit of logistic regression models was assessed with the C-statistic (which is the area under the ROC curve, and is used as a measure of discrimination) presented with 95% confidence intervals (CI). To assess the value added by including the MR-proADM level with the NEWS score in predicting deterioration, continuous net reclassification improvement (NRI) and integrated discrimination improvement (IDI) were calculated[32 33].

For internal validation of the statistical models the C-statistic was evaluated after correcting for optimistic predictions through bootstrapping with 10,000 resamples.

# Results

## Study enrolment

The process of recruitment and enrolment of patients for the study is shown in Figure 1. The study recruited 300 patients, and 292 were included in the analysis. Five patients were excluded because the blood samples for MR-proADM were taken more than 12 hours from baseline NEWS assessment; 3 patients were excluded from the primary outcome due to missing follow up NEWS scores.

## Patient characteristics

Patient demographics and mean biomarker levels for each covariate are reported in Table 1. The cohort was evenly divided in gender and had a mean age of 63 years and mean NEWS on admission of 3, with the majority of patients having NEWS score of 2. COPD or HF were present in 28%, and 25% had other comorbidities.

Prognostic accuracy of MR-proADM in emergency departments

**Table 1**. Characteristics of the study population, classified by Outcome 1 (*Acuity Increase*), Outcome 2 (*Deterioration Event*) and All patients. Data are presented as number (no) and percentages (%) for counts, or mean and (standard deviation, SD) for continuous normally distributed data, or [25th; 50th; 75th percentile] for continuous non-normally distributed data.

| | **Outcome 1: *Acuity Increase*** | | **Outcome 2: *Deterioration Event*** | | **All patients (n = 292)** |
|---|---|---|---|---|---|
| | **Present (e = 84)** | **Absent** | **Present (e2 = 32)** | **Absent** | |
| Age (mean years, SD) | 65 (17) | 62 (21) | 63 (14) | 63 (20) | 63 (20) |
| Gender (no. females, %) | 41 (49%) | 107 (51%) | 15 (47%) | 133 (51%) | 148 (51%) |
| NEWS = 2 (no., %) | 34 (40%) | 82 (40%) | 12 (38%) | 104 (40%) | 116 (40%) |
| NEWS = 3 (no., %) | 26 (31%) | 59 (28%) | 9 (28%) | 76 (29%) | 85 (29%) |
| NEWS = 4 (no., %) | 11 (13%) | 43 (21%) | 4 (13%) | 50 (19%) | 54 (18%) |
| NEWS = 5 (no., %) | 13 (15%) | 24 (12%) | 7 (22%) | 30 (12%) | 37 (13%) |
| MR-proADM (mean nmol/l, SD) | 1.50 (1.4)<br>[0.72, 1.12, 1.79] | 1.19 (0.9)<br>[0.68, 0.93, 1.28] | 1.89 (2.0)<br>[0.93, 1.13, 1.95] | 1.20 (0.9)<br>[0.68, 0.93, 1.39] | 1.28 (1.1)<br>[0.68, 0.97, 1.48] |
| CRP (mg/l) | 59 (79)<br>[5, 22, 80] | 42 (70)<br>[4, 13, 41] | 61 (90)<br>[7, 23, 67] | 45 (71)<br>[4, 16, 51] | 47 (73)<br>[4, 17, 54] |
| WBC (x$10^9$/l) | 12 (5)<br>[9, 10, 14] | 11 (5)<br>[8, 10, 14] | 12 (4)<br>[9, 12, 15] | 11 (5)<br>[8, 10, 14] | 11 (5)<br>[8, 10, 14] |
| COPD/HF (no, %)* | 33 (39%) | 46 (22%) | 12 (38%) | 67 (26%) | 79 (28%) |
| Other comorbidities (no., %) | 17 (20%) | 55 (26%) | 15 (47%) | 57 (22%) | 72 (25%) |
| *Length of Stay* (hrs) | 168 (196)<br>[63, 110, 194] | 137 (176)<br>[26, 68, 176] | 173 (172)<br>[59, 106, 259] | 143 (172)<br>[33, 72, 176] | 146 (182)<br>[35, 77, 182] |
| Length of Stay in MAU (hrs) | 31 (19)<br>[17, 25, 43] | 24 (16)<br>[13, 21, 30] | 27 (17)<br>[18, 23, 35] | 26 (17)<br>[15, 22, 31] | 26 (17)<br>[15, 22, 31] |
| Monitored beds (no, %) | 31 (37%) | 58 (27%) | 11 (34%) | 78 (30%) | 89 (30%) |

Prognostic accuracy of MR-proADM in emergency departments

| | Outcome 1: *Acuity Increase* | | Outcome 2: *Deterioration Event* | | All patients (n = 292) |
|---|---|---|---|---|---|
| | Present (e = 84) | Absent | Present (e2 = 32) | Absent | |
| Deterioration time (hrs) | 15 (13) [5, 9, 21] | N/A | 170 (226) [19, 33, 301] | N/A | |
| * for COPD: e = number with *Acuity Increase* = 82; e2 = number with *Deterioration Event = 29;* n = total number of patients = 282 | | | | | |

221

**Table 2**. Criteria met by patients classified with an *Acuity Increase* or *Deterioration Event.*

| Criterion for deterioration | *Acuity Increase* (e = 84) | *Deterioration Event* (e2 = 32) |
|---|---|---|
| NEWS (no, %) | 81 (96.4%) | N/A |
| ICU transfer (no, %) | 1 (1.2%) | 4 (12.5%) |
| Death (no, %) | 0 (0%) | 6 (18.8%) |
| Readmission (no, %) | 2 (2.4%) | 22 (68.7%) |

The study population was homogenous across *Acuity Increase* and No *Acuity Increase* outcomes in terms of gender, age, and NEWS on admission. Table 2 shows the frequencies of criteria determining *Acuity Increase* and *Deterioration Event*. Notably, around 95% of *Acuity Increases* were the result of an increase in NEWS score, while readmission was the reason for around 70% of *Deterioration Event*s.

Patients who experienced *Acuity Increase* had higher MR-proADM and CRP levels at admission, and longer *Length of Stay* in the hospital and in the MAU.

The prevalence of *Acuity Increase* was 29% (somewhat higher than the anticipated 20%). The prevalence of *Deterioration Event*s was 11%. The numbers of events provided sufficient statistical power to assess statistical significance for the primary outcome, *Acuity Increase*, but not for the secondary outcome, *Deterioration Event*.

## Accuracy of MR-proADM for predicting *Acuity Increase*

Potentially useful predictors with univariate analysis of *Acuity Increase* were MR-proADM (OR = 1.27, 95% CI 1.02, 1.62; p = 0.037), $Age^2$ (OR = 1.00, 95% CI 0.99, 1.00; p = 0.023) and the presence of COPD or HF (OR = 2.25, 95% CI1.30, 3.91; p = 0.004; Supplementary Figure s1). The prognostic accuracy of CRP, WBC and NEWS did not reach the threshold of significance (p = 0.88, p = 0.090, Table 3, and p=0.416, Table 4, respectively).

Prognostic accuracy of MR-proADM in emergency departments

**Table 3.** Univariate regression analyses for predicting the three outcomes of interest: *Acuity Increase*, *Deterioration Event*, and *Length of Stay*. The p-values are for the statistical significance of the corresponding covariate in the related model. Analyses for the NEWS score as a predictor are shown in Table 4.

| | Beta | CI | Odds Ratio (CI) | p-value |
|---|---|---|---|---|
| ***Acuity Increase*: univariate logistic regressions (n = 292, e = 84)** | | | | |
| MR-proADM | 0.24 | -0.02, 0.48 | 1.27 (1.02, 1.62) | 0.037 |
| CRP | 0.003 | -0.0005, 0.0063 | 1.00 (1.00, 1.01) | 0.088 |
| WBC | 0.04 | -0.008, 0.094 | 1.05 (1.00, 1.10) | 0.09 |
| Gender | 0.14 | -0.38, 0.65 | 1.15 (0.69, 1.92) | 0.684 |
| Age | 0.1 | 0.019, 0.1925 | 1.11 (1.02, 1.21) | 0.023 |
| $Age^2$ | -0.0008 | -0.0016, -0.0001 | 1.00 (0.99, 1.00) | |
| Other Comorbidities | -0.32 | -0.96, 0.28 | 0.72 (0.38, 1.32) | 0.267 |
| COPD/HF* | 0.81 | 0.26, 1.36 | 2.25 (1.30, 3.91) | 0.004 |
| ***Deterioration Event*: univariate logistic regressions (n = 292, e2 = 32)** | | | | |
| MR-proADM | 0.37 | 0.11, 0.64 | 1.44 (1.12, 1.90) | 0.006 |
| CRP | 0.003 | -0.002, 0.01 | 1.00 (1.00, 1.01) | 0.255 |
| WBC | 0.02 | -0.05, 0.09 | 1.02 (0.95, 1.10) | 0.506 |
| Gender | 0.17 | -0.57, 0.92 | 1.19 (0.57, 2.50) | 0.648 |
| Age | 0.21 | 0.06, 0.40 | 1.23 (1.06, 1.49) | 0.013 |
| $Age^2$ | -0.002 | -0.003, -0.001 | 1.00 (1.00, 1.00) | |
| Other Comorbidities | 1.14 | 0.38, 1.90 | 3.14 (1.47, 6.69) | 0.003 |
| COPD/HF* | 0.67 | -0.14, 1.46 | 1.96 (0.87, 4.29) | 0.095 |
| ***Length of Stay*: simple linear regressions (n = 292, e = 84, e2 = 32 )** | | | | |
| MR-proADM | 0.7 | 0.49, 0.92 | N/A | <0.0001 |
| CRP | 0.05 | -0.05, 0.15 | N/A | 0.368 |
| WBC | -0.06 | -0.38, 0.27 | N/A | 0.73 |
| Gender | 0.08 | -0.04, 0.20 | N/A | 0.18 |
| Age | 0.007 | 0.004, 0.010 | N/A | <0.0001 |
| Other Comorbidities | 0.18 | 0.05, 0.32 | N/A | 0.009 |
| COPD/HF* | 0.07 | -0.07, 0.21 | N/A | 0.318 |

**Key**: n = total number of cases; e = number of *Acuity Increases*; $e_2$ = number of *Deterioration Events*; CI = 95% confidence interval
* n = 282, e = 82, e2 = 29

**Table 4.** Multivariable regression analyses for the outcomes of interest: *Acuity Increase*, *Deterioration Event*, *Length of Stay* (Outcomes 1, 2, and 3 respectively) with NEWS comparator group. Predictor set *a.* includes only the NEWS score as a predictor; Predictor set *b.* includes MR-proADM and NEWS scores; Predictor set *c*. includes MR-proADM, NEWS scores, and other significant predictors and interactions. . The p-values are for the statistical significance of the corresponding covariate in the related model.

| | | Beta | CI | Odds Ratio (CI) | p-value |
|---|---|---|---|---|---|
| ***Acuity Increase*: multivariate logistic regressions** | | | | | |
| **Predictor set *a*** n = 292 e = 84 e2 = 32 | NEWS 3 | 0.06 | -0.55, 0.67 | 1.06 (0.57, 1.95) | 0.416 |
| | NEWS 4 | -0.48 | -1.29, 0.27 | 0.62 (0.27, 1.31) | |
| | NEWS 5 | 0.27 | -0.54, 1.04 | 1.31 (0.58, 2.84) | |
| **Predictor set *b*** n = 292 e = 84 e2 = 32 | NEWS 3 | 0.03 | -0.59, 0.65 | 1.03 (0.56, 1.91) | 0.247 |
| | NEWS 4 | -0.53 | -1.35, 0.23 | 0.59 (0.26, 1.26) | |
| | NEWS 5 | 0.18 | -0.63, 0.97 | 1.20 (0.53, 2.64) | |
| | MR-proADM | 0.24 | 0.02, 0.49 | 1.28 (1.02, 1.63) | 0.039 |
| **Predictor set *c*** n = 282 e = 82 e2 = 29 | NEWS 3 | -0.11 | -0.76, 0.54 | 0.90 (0.47, 1.71) | 0.221 |
| | NEWS 4 | -0.89 | -1.77, -0.08 | 0.41 (0.17, 0.93) | |
| | NEWS 5 | 0.09 | -0.77, 0.91 | 1.09 (0.46, 2.50) | |
| | MR-proADM | 0.41 | 0.13, 0.76 | 1.51 (1.14, 2.14) | 0.01 |
| | COPD/HF | 1.81 | 0.80, 2.85 | 6.08 (2.23, 17.35) | 0.001 |
| | MR-proADM*COPD/HF | -0.71 | -1.40, -0.10 | 0.49 (0.25, 0.91) | 0.03 |
| ***Deterioration Event*: multivariate logistic regressions** | | | | | |
| **Predictor set *a*** n = 292 e = 84 e2 = 32 | NEWS 3 | 0.03 | -0.92, 0.94 | 1.03 (0.40, 2.55) | 0.512 |
| | NEWS 4 | -0.37 | -1.68, 0.74 | 0.69 (0.19, 2.10) | |
| | NEWS 5 | 0.7 | -0.36, 1.70 | 2.02 (0.70, 5.50) | |
| **Predictor set *b*** *n* = 292 e = 84 e2 = 32 | NEWS 3 | -0.01 | -0.97, 0.92 | 0.99 (0.38, 2.51) | 0.564 |
| | NEWS 4 | -0.43 | -1.76, 0.70 | 0.65 (0.17, 2.02) | |
| | NEWS 5 | 0.6 | -0.49, 1.62 | 1.81 (0.61, 5.05) | |
| | MR-proADM | 0.36 | 0.10, 0.64 | 1.43 (1.11, 1.89) | 0.007 |

Prognostic accuracy of MR-proADM in emergency departments

| | | Beta | CI | Odds Ratio (CI) | p-value |
|---|---|---|---|---|---|
| **Predictor set *c*** n = 282 e = 82 e2 = 29 | NEWS 3 | 0.16 | -0.83, 1.12 | 1.17 (0.44, 3.07) | 0.389 |
| | NEWS 4 | -0.49 | -1.86, 0.69 | 0.62 (0.16, 2.00) | |
| | NEWS 5 | 0.69 | -0.44, 1.76 | 1.99 (0.64, 5.81) | |
| | MR-proADM | 0.32 | 0.02, 0.64 | 1.37 (1.02, 1.89) | 0.044 |
| | Other comorbidities | 0.94 | 0.10, 1.77 | 2.56 (1.10, 5.85) | 0.026 |
| | Age | 0.21 | 0.06, 0.41 | 1.23 (1.06, 1.50) | 0.011 |
| | $Age^2$ | -0.002 | -0.003, -0.001 | 1.00 (1.00, 1.00) | |
| ***Length of Stay*: multiple linear regressions** | | | | | |
| **Predictor set *a*** n = 292 e = 84 e2 = 32 | NEWS 3 | -0.07 | -0.21, 0.08 | N/A | 0.052 |
| | NEWS 4 | 0.07 | -0.10, 0.24 | N/A | |
| | NEWS 5 | 0.21 | 0.01, 0.40 | N/A | |
| **Predictor set *b*** *n = 292* *e = 84* *e2 = 32* | NEWS 3 | -0.1 | -0.24, 0.04 | N/A | 0.033 |
| | NEWS 4 | 0.05 | -0.11, 0.21 | N/A | |
| | NEWS 5 | 0.14 | -0.04, 0.32 | N/A | |
| | MR-proADM | 0.69 | 0.48, 0.91 | N/A | <0.0001 |
| **Predictor set *c*** n = 282 e = 82 e2 = 29 | NEWS 3 | -0.12 | -0.25, 0.02 | N/A | 0.031 |
| | NEWS 4 | 0.04 | -0.11, 0.20 | N/A | |
| | NEWS 5 | 0.14 | -0.04, 0.32 | N/A | |
| | MR-proADM | 0.55 | 0.31, 0.80 | N/A | <0.0001 |
| | Age | 0.004 | 0, 0.007 | N/A | 0.027 |

The prognostic accuracy for *Acuity Increase* of NEWS on its own was limited and not significant (AUC 0.55, 95% CI 0.48, 0.62), but when MR-proADM was included as an additional predictor, the accuracy of the model increased (AUC 0.61, 95% CI 0.54, 0.69; OR = 1.28, 95% CI 1.02, 1.63; p = 0.039) (Tables 4 and 5, Figure 2 panel A). When including MR-proADM with NEWS, the reclassification of patients was also significant, especially for the NRI (NRI = 0.3, SE 0.1, p = 0.007; IDI = 0.017, Table 4).

The prognostic accuracy of MR-proADM and the additional value it provides to the NEWS score was confirmed for *Deterioration Event*s and *Length of Stay* (Tables 4 and 5, and Figure 2 panels B and C).

For MR-proADM alone, the AUCs were for: *Acuity Increase* 0.58 (0.51-0.66), and *Deterioration Event* 0.64 (0.54-0.74). For *Length of Stay* the R squared was 0.12.

**Table 5.** Model comparisons. Outcomes 1, 2, and 3 refer to *Acuity Increase*, *Deterioration Event*, and *Length of Stay* respectively. The predictors are: *Set a* NEWS score alone; *Set b* NEWS score and MR-proADM; *Set c* NEWS score, MR-proADM, and other significant predictors and interactions detailed in Table 3.

| | AIC | Deviance | AUC (CI) or $R^2$ for linear regression | NRI (se), p-value | IDI (se), p-value |
|---|---|---|---|---|---|
| Outcome 1 - predictor set *a.* | 348 | 356 | 0.55 (0.48, 0.62) | | |
| Outcome 1 - predictor set *b.* | 343 | 353 | 0.61 (0.54, 0.69) | 0.3 (0.1), 0.007 | 0.017 (0.009), 0.058 |
| Outcome 1 - predictor set *c.* | 317 | 331 | 0.69 (0.63, 0.76) | 0.4 (0.1), 0.0004* | 0.05 (0.01),0.0009* |
| Outcome 2 - predictor set *a.* | 199 | 207 | 0.57 (0.47, 0.68) | | |
| Outcome 2 - predictor set *b.* | 192 | 202 | 0.65 (0.54, 0.76) | 0.4 (0.2), 0.003 | 0.04 (0.02), 0.10 |
| Outcome 2 - predictor set *c.* | 177 | 193 | 0.73 (0.63, 0.84) | 0.5 (0.2), 0.012* | 0.06 (0.02), 0.0004* |
| Outcome 3 - predictor set *a.* | 77 | -381 | 0.03 | | |
| Outcome 3 - predictor set *b.* | 68 | -417 | 0.14 | | |
| Outcome 3 - predictor set *c.* | 67 | -420 | 0.16 | | |

## Effect on prognostic accuracy when clinical information is added to the set of predictors

Secondary multivariable modelling evaluated the prognostic accuracy of MR-proADM when adjusted for the clinical factors in predictive set *c*: age, gender, CRP, WBC, presence of COPD or HF, presence of other comorbidities,

For *Acuity Increase,* COPD or HF comorbidity status and its interaction with MR-proADM level improved the prognostic accuracy of the model: AUC increased from 0.61 (95% CI 0.54, 0.69) to 0.69 (95% CI 0.63, 0.76), and net reclassification index from 0.3 to 0.4 (Table 5).

For *Deterioration Event*s, the presence of other comorbidities (excluding COPD and HF) and $Age^2$ increased the prognostic accuracy of MR-proADM, (Table 4 and 5). The

prognostic accuracy of *Length of Stay* (Outcome 3) of MR-proADM is also increased by including *Age* in the model (Table 4 and 5, Supplementary Figure s2).

## Potential confounding effects

**Shorter term outcomes**: NEWS and MR-proADM were less accurate in predicting *Acuity Increase* within 24 and 12 hours from admission than in predicting *Acuity Increase* within 72 hours (Supplementary Tables s1 and s2).

**Interval between admission NEWS scoring and blood collection:** Because ward processes did not allow the times of scoring NEWS and collecting blood to be specified for research, we assessed for a confounding effect from variation in the timings, but found no evidence for it (Supplementary Table s3).

**Correlations among biomarkers**. Diagnostic plots, shown in Supplementary Figures s2 and s3, show no multicollinearity in the data, no autocorrelation, no heteroscedasticity, and no data points that stood out in terms of their influence on results.

## Sensitivity and specificity

As overall measures of accuracy, sensitivity and specificity were calculated (where appropriate) for each model using Youden's index. The results are shown in Supplementary Table s4. In practice, the trade-off between sensitivity and specificity would depend on the type of clinical decision to be made on the result (i.e. "rule-in" or "rule out") and this would differ from the approach in Youden's Index, which gives equal weight to false positive and false negative results.

## Internal Validation

C-statistic values after correcting for optimistic predictions (i.e. bootstrapped average of the AUC for each model) were: for *Acuity Increase:* predictor set *a*, C-stat=0.53*;* predictor set *b*, C-stat=0.59; predictor set *c*, C-stat=0.66. For *Deterioration Events:* predictor set *a*, C-stat=0.52; predictor set *b*, C-stat=0.61, predictor set *c*, C-stat=0.68. For *Length of Stay:* predictor set *a*, $R^2$=0.003; predictor set *b*, $R^2$=0.12; predictor set *c*, $R^2$=0.13. AUCs decreased slightly with the bootstrapped averages, but the differences between the AUCs for Predictor sets *a, b*, and *c* were constant. These results are an internal validation, and further validation on an external dataset is required.

# Discussion

## Accuracy of prediction of deterioration by MR-proADM

This study shows that MR-proADM may be a clinically useful biomarker for predicting deterioration (i.e. *Acuity Increase*) within 72 hours from admission to hospital in mild to moderately ill patients with admission NEWS score between 2 to 5. By design, NEWS scores in this range imply a low risk of deterioration, and our data are consistent with this. Previous evaluations of the NEWS score assessed on admission have found that it predicts deterioration [3-5 34], which may seem inconsistent. But these studies included all patients admitted to ED, whatever their NEWS score.

For most of the observed *Acuity Increase* events, the reason for classification was an increase in the NEWS score. Because an increase in NEWS score reflects both measurement variation and physiological variation, additional exploratory analyses were carried out to assess the performance of MR-proADM, using an operational definition of deterioration, *Deterioration Event*, designed to minimize measurement variation. NEWS on its own had low prognostic accuracy for *Deterioration Event*s. However, MR-proADM level, and NEWS score together predicted *Deterioration Events* with an AUC of 0.65. Considering baseline patient characteristics further increased the accuracy of the model (AUC = 0.73).

## Comorbidities and interactions with MR-proADM levels

MR-proADM levels in people with COPD and/or heart failure are chronically raised and are not predictive of deterioration. However, in other people whose MR-proADM levels are not chronically raised, high levels are predictive of *Acuity Increase* (Supplementary Figure s1). Including these comorbidities and their interaction with MR-proADM level increased the prognostic accuracy of the logistic regression model.

## Limitations

This study included only patients who were admitted with a NEWS score between 2 and 5. The prognostic accuracy of the MR-proADM would perhaps have been greater if more extreme cases had been included. However, patients with NEWS scores more than 5 are already known to be severely ill and to require close monitoring and/or management at higher levels of care.

Internal validation found that the uncorrected C-statistics are optimistic, which implies that external validation in an independent study would be useful. However, after correction for optimistic predictions, the study's conclusions remain unchanged.

## Interpretations and implications

The contributions of MR-proADM to the accuracy of the prognostic models suggests that it could provide additional prognostic information over and above NEWS score.

Secondary analyses suggest that a potentially useful clinical decision aid could be based on the NEWS score, MR-proADM level, and clinical features.

## Future research and development

As a growing number of NHS hospitals are implementing the NEWS score on their clinical information systems, it should be practical to develop a decision aid based on admission NEWS score, MR-proADM level, and clinical features. Other biomarkers may further improve prognostic accuracy for deterioration, for example: lactate [3]; peroxiredoxin-4 (Prx4) and copeptin[22 35 36]; and soluble urokinase plasminogen activator receptor (suPAR)[37]. The feasibility, cost-effectiveness, and acceptability of such decision aids needs to be evaluated in further research.

A rapid point of care test for MR-proADM could facilitate the assessment process and reduce delays in arranging optimal levels of care and intensity of monitoring. Future research could identify the threshold MR-proADM level corresponding to the optimal combination of sensitivity and specificity for a binary test (e.g. "present" or "absent") for deterioration.

# Footnotes

**Contributors:** AJS, DAP, MP, and DS devised the study; RO, SG, and AJA managed the project; SG performed the statistical analyses with advice from DS; all authors contributed to the final manuscript.

**Acknowledgements**: The authors would like to thank the patients who accepted to be part of the study, and the team of research nurses for the data collection: Laura Shewan, Louise Taylor, Janine Gradwell, Karen Martin, Katherine Cullen, Gerry Jones, Graham Soulsby, Carmen Scott, and Helen Reed. We also thank the reviewers for their helpful

suggestions and comments. The views expressed are those of the authors and not necessarily those of B·R·A·H·M·S GmbH, the NHS, the NIHR, or the Department of Health.

**Funding:** The study was funded by B·R·A·H·M·S GmbH. The NIHR funds the NIHR Newcastle In Vitro Diagnostics Co-operative.

**Competing interests**: None declared.

**Ethics and governance approvals:** The study was approved by the Newcastle & North Tyneside Research Ethics Committee (15/NE/0120), and by the R&D Committee of the Newcastle upon Tyne Hospitals NHS foundation Trust (reference number 7495).

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data are available.

References

1. Royal College of Physicians of London. National early warning score (NEWS): standardising the assessment of acute-illness severity in the NHS—report of a working party. . London, 2012.
2. McGinley A, Pearse RM. A national early warning score for acutely ill patients. BMJ : British Medical Journal 2012;**345**.
3. Jo S, Yoon J, Lee JB, et al. Predictive value of the National Early Warning Score–Lactate for mortality and the need for critical care among general emergency department patients. Journal of Critical Care 2016;**36**:60-68.
4. Smith GB, Prytherch DR, Meredith P, et al. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation 2013;**84**(4):465-70.
5. Smith ME, Chiovaro JC, O'Neil M, et al. Early warning system scores for clinical deterioration in hospitalized patients: a systematic review. Annals of the American Thoracic Society 2014;**11**(9):1454-65.
6. Teasdale GM. National early warning score (NEWS) is not suitable for all patients. BMJ : British Medical Journal 2012;**345**.
7. Eccles SR, Subbe C, Hancock D, et al. CREWS: Improving specificity whilst maintaining sensitivity of the National Early Warning Score in patients with chronic hypoxaemia. Resuscitation 2014;**85**(1):109-11.
8. Badriyah T, Briggs JS, Meredith P, et al. Decision-tree early warning score (DTEWS) validates the design of the National Early Warning Score (NEWS). Resuscitation 2014;**85**(3):418-23.
9. Abbott TE, Vaid N, Ip D, et al. A single-centre observational cohort study of admission National Early Warning Score (NEWS). Resuscitation 2015;**92**:89-93.
10. Corfield AR, Lees F, Zealley I, et al. Utility of a single early warning score in patients with sepsis in the emergency department. Emergency medicine journal : EMJ 2014;**31**(6):482-7.

11. Bilben B, Grandal L, Søvik S. National Early Warning Score (NEWS) as an emergency department predictor of disease severity and 90-day survival in the acutely dyspneic patient – a prospective observational study. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine 2016;**24**:80.

12. Valenzuela-Sanchez F, Valenzuela-Mendez B, Rodriguez-Gutierrez JF, et al. New role of biomarkers: mid-regional pro-adrenomedullin, the biomarker of organ failure. Annals of translational medicine 2016;**4**(17):329.

13. Stokes NR, Dietz BW, Liang JJ. Cardiopulmonary laboratory biomarkers in the evaluation of acute dyspnea. Open access emergency medicine : OAEM 2016;**8**:35-45.

14. Sinning C, Zengin E, Zeller T, et al. Candidate biomarkers in heart failure with reduced and preserved ejection fraction. Biomarkers : biochemical indicators of exposure, response, and susceptibility to chemicals 2015;**20**(4):258-65.

15. Kutz A, Hausfater P, Amin D, et al. The TRIAGE-ProADM Score for an Early Risk Stratification of Medical Patients in the Emergency Department - Development Based on a Multi-National, Prospective, Observational Study. PloS one 2016;**11**(12).

16. Albrich WC, Dusemund F, Ruegger K, et al. Enhancement of CURB65 score with proadrenomedullin (CURB65-A) for outcome prediction in lower respiratory tract infections: derivation of a clinical algorithm. BMC infectious diseases 2011;**11**:112.

17. Lopez J, Martinez A. Cell and molecular biology of the multifunctional peptide, adrenomedullin. International review of cytology 2002;**221**:1-92.

18. Serrano-Ponz M, Rodrigo-Gasque C, Siles E, et al. Temporal profiles of blood pressure, circulating nitric oxide, and adrenomedullin as predictors of clinical outcome in acute ischemic stroke patients. Molecular medicine reports 2016;**13**(5):3724-34.

19. Zudaire E, Portal-Nunez S, Cuttitta F. The central role of adrenomedullin in host defense. Journal of leukocyte biology 2006;**80**(2):237-44.

20. Zuur-Telgen MC, Brusse-Keizer MG, Vandervalk PD, et al. Stable state MR-proadrenomedullin level is a strong predictor for mortality in COPD patients. Chest 2013.

21. Pedowska-Wloszek J, Kostrubiec M, Kurnicka K, et al. Midregional proadrenomedullin (MR-proADM) in the risk stratification of patients with acute pulmonary embolism. Thrombosis research 2013;**132**(5):506-10.

22. Nickel CH, Messmer AS, Geigy N, et al. Stress markers predict mortality in patients with nonspecific complaints presenting to the emergency department and may be a useful risk stratification tool to support disposition planning. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2013;**20**(7):670-9.

23. Landman GW, van Dijk PR, Drion I, et al. Midregional fragment of proadrenomedullin, new-onset albuminuria, and cardiovascular and all-cause mortality in patients with type 2 diabetes (ZODIAC-30). Diabetes care 2014;**37**(3):839-45.

24. Funke-Kaiser A, Mann K, Colquhoun D, et al. Midregional proadrenomedullin and its change predicts recurrent major coronary events and heart failure in stable coronary heart disease patients: the LIPID study. International journal of cardiology 2014;**172**(2):411-8.

25. Gaggin HK, Januzzi JL, Jr. Biomarkers and diagnostics in heart failure. Biochimica et biophysica acta 2013;**1832**(12):2442-50.

26. Vigue B, Leblanc PE, Moati F, et al. Mid-regional pro-adrenomedullin (MR-proADM), a marker of positive fluid balance in critically ill patients: results of the ENVOL study. Critical care (London, England) 2016;**20**(1):363.

27. Tzikas S, Keller T, Ojeda FM, et al. MR-proANP and MR-proADM for risk stratification of patients with acute chest pain. Heart (British Cardiac Society) 2013;**99**(6):388-95.

28. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. Journal of clinical epidemiology 1996;**49**(12):1373-9.

29. van Smeden M, de Groot JAH, Moons KGM, et al. No rationale for 1 variable per 10 events criterion for binary logistic regression analysis. BMC Medical Research Methodology 2016;**16**(1):163.

30. Travaglino F, Russo V, De Berardinis B, et al. Thirty and ninety days mortality predictive value of admission and in-hospital procalcitonin and mid-regional pro-adrenomedullin testing in patients with dyspnea. Results from the VERyfing DYspnea trial. The American journal of emergency medicine 2014;**32**(4):334-41.

31. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria.: R Foundation for Statistical Computing, 2011.

32. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., et al. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Statistics in medicine 2008;**27**(2):157-72; discussion 207-12.

33. Steyerberg EW, Vedder MM, Leening MJ, et al. Graphical assessment of incremental value of novel markers in prediction models: From statistical to decision analytical perspectives. Biometrical journal Biometrische Zeitschrift 2015;**57**(4):556-70.

34. Alam N, Vegting IL, Houben E, et al. Exploring the performance of the National Early Warning Score (NEWS) in a European emergency department. Resuscitation 2015;**90**:111-5.

35. Nickel CH, Ruedinger J, Misch F, et al. Copeptin and peroxiredoxin-4 independently predict mortality in patients with nonspecific complaints presenting to the emergency department. Academic emergency medicine : official journal of the Society for Academic Emergency Medicine 2011;**18**(8):851-9.

36. Iversen K, Gotze JP, Dalsgaard M, et al. Risk stratification in emergency patients by copeptin. BMC medicine 2014;**12**:80.

37. Rasmussen LJ, Ladelund S, Haupt TH, et al. Soluble urokinase plasminogen activator receptor (suPAR) in acute care: a strong marker of disease presence and severity, readmission and mortality. A retrospective cohort study. Emergency medicine journal : EMJ 2016.

# Figure legends

**Figure 1.** Patient recruitment process.

**Figure 2. Panel A.** Prognostic accuracy for *Acuity Increase*; predictor set *a*: NEWS; predictor set *b*: NEWS, MR-proADM; predictor set *c*: NEWS, MR-proADM, COPD/HF, interaction between MR-proADM and COPD/HF. **Panel B.** Comparisons as for panel **A** but for predicting a *Deterioration Event*; predictor set *c:* NEWS score, MR-proADM level, Age$^2$, other comorbidities. **Panel C.** *Length of Stay* predicted by MR-proADM level.

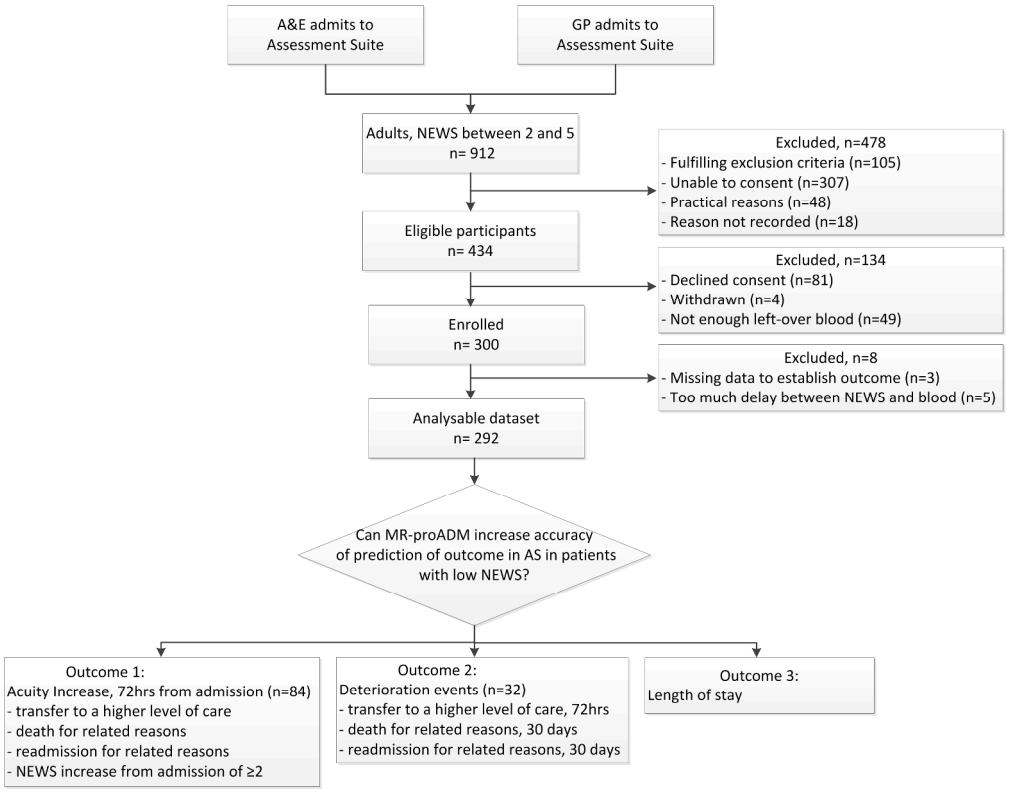



Figure 1. Patient recruitment process.

255x200mm (300 x 300 DPI)

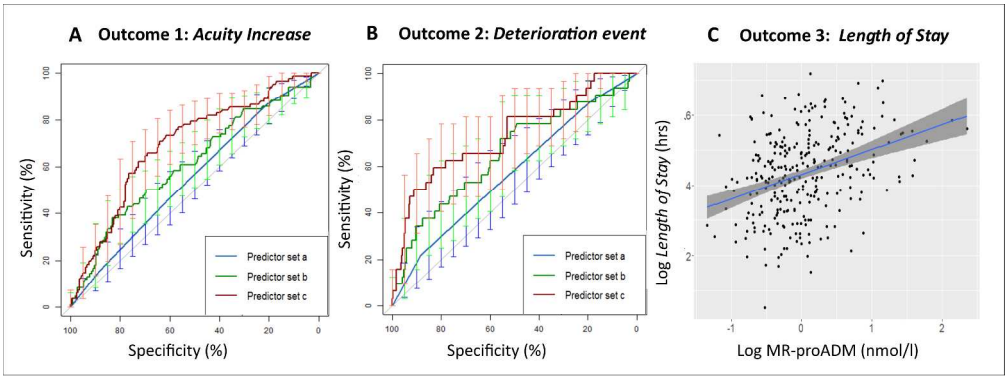



Figure 2. Panel A. Prognostic accuracy for Acuity Increase; predictor set a: NEWS; predictors set b: NEWS, MR-proADM; predictor set c: NEWS, MR-proADM, COPD/HF, interaction between MR-proADM and COPD/HF. Panel B. Comparisons as for panel A but for predicting a Deterioration Event; predictor set c: NEWS score, MR-proADM level, Age2, other comorbidities. Panel C. Length of Stay predicted by MR-proADM level.

Link text : Figure 2

320x118mm (300 x 300 DPI)

Supplementary information for Graziadio et al, 2018

Supplementary information for *Can mid-regional pro-adrenomedullin (MR-proADM) increase the prognostic accuracy of NEWS in predicting deterioration in patients admitted to hospital with mild to moderately severe illness? A prospective single-centre observational study*. Graziadio et al, 2018

## Table of Contents

Supplementary information for Graziadio et al, 2018

# Additional information on Methods

## NEWS as a predictor of deterioration in patients with mild to moderately severe illness

### Operational definitions of deterioration

In the original validation of NEWS, about 2% of the population had the combined outcome of cardiac arrest, unanticipated ICU admission, or death — each within 24 hours [Smith 2013]. Furthermore, the proportions for each of the three individual outcomes and the composite outcomes increased monotonically through the range of NEWS scores.

Thus, as designed in its development, a NEWS score between 2 and 5 defines a population at low risk of cardiac arrest within 24 hours, death within 24 hours, or ICU admission within 24 hours.

However, a clinically important proportion of patients admitted to A&E or Medical Admissions Unit with mild to moderately severe illness (NEWS scores between 2 and 5) do deteriorate, and the NEWS score is, by design, not able to identify these patients.

The improvement challenge is to identify biomarkers that will increase the discrimination of low NEWS scores. And the methodological challenge was to develop convenient and effective operational definitions of deterioration from mild/moderately severe.

As NEWS is used to monitor changes in severity of illness, we decided to base our primary operational definition of deterioration on an increase of at least 2 in the NEWS score.

***Acuity Increase***. The primary outcome was the proportion of patients who, within 72 hours, had any combination of:

- an increase of at least 2 in the NEWS score
- transfer to a higher-dependency bed or monitored area
- death
- for those discharged from hospital, re-admission for medical reasons.

We labelled this measure *Acuity Increase*.

Because there was concern about variations in NEWS scoring and about using NEWS to predict a change in NEWS (which it is designed not to do in this study's population), we defined two other measures of deterioration, one direct, and the other indirect*:*

***Deterioration Event***: the occurrence of one or more of the following:

- transfer to higher level of care within 72 hours from admission;
- death (for reasons related to admission) within 30 days;
- re-admission to hospital (for the same reason as the previous admission) within 30 days from first admission

***Length of Stay***: the duration in days from admission to discharge or death

Supplementary information for Graziadio et al, 2018

### Predictors of deterioration for statistical modelling

In the analyses, we included NEWS as a possible predictor of deterioration. As expected, NEWS scores consistently do not predict deterioration, for all three of our operational definitions.

In line with the original validation of NEWS, we included NEWS as an ordinal variable [Smith 2013].

## Analytical data exploration

Univariate logistic regressions were used to investigate whether the relationship between outcome variables (i.e. deterioration measures) and the input variables (NEWS and MR-proADM, age, comorbidities, gender, CRP, and WBC) were linear. If they were not linear, log transformation and squared transformation were applied. If the transformation substantially lowered the AIC, then the transformed variable was used in statistical analyses.

For categorical variables with multiple ordinal levels (i.e. NEWS score), the univariate analysis informed if it was appropriate to include the variable in the model as a continuous or categorical factor. If the coefficients in the univariate models increased linearly, then a linear relationship with the outcome could be assumed, and the variable was included in the model as continuous, otherwise the variable was treated as a categorical factor.

Univariate analyses were also used to identify the variables that affected the outcome significantly. Variables with a probable relationship with the outcome variable ($p<0.1$) were included in the full model logistic regression.

## Visual data exploration and interaction between MR-proADM and COPD/HF

After formatting the datasets, all variables were graphed (bar-charts for categorical variables, and scatterplots/histograms for continuous variables) and visually checked for outliers and distributions that seemed potentially erroneous.

If outliers were identified, the cause(s) were investigated to understand whether they were due to human error or they were genuine data. Outliers were kept in the primary analysis. In a secondary sensitivity analysis, outliers were removed and the same analyses repeated to assess the impact on the results. If the coefficients of the predictors changed substantially, both models would be described. There was one genuine outliner patient with a very high level of MR-proADM compared to the population mean, but its exclusion made no meaningful difference to the results, and the subject was included in the final analysis.

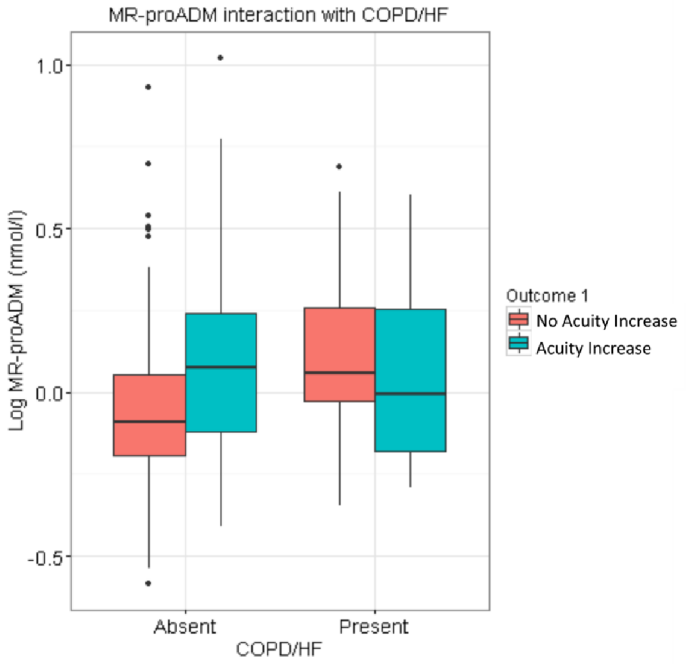
The influence of potentially important factors (such as comorbidities) on the ability of the MR-proADM to predict deterioration was explored graphically.

### Interaction between MR-proADM and COPD/HF

A significant interaction between MR-proADM and the presence of COPD/HF was discovered, and therefore included in the logistic regression (Outcome 1, predictor set c). The plot is

shown in **Supplementary Figure s1**. This interaction showed that the MR-proADM level was increased in patients who deteriorated, but only if they did not have COPD or HF.

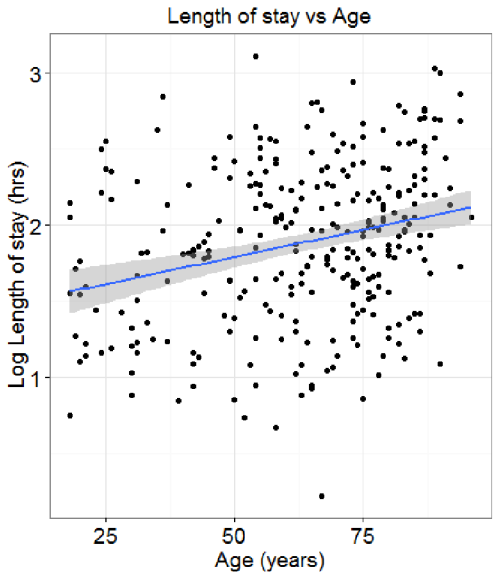
There was no suggestion that age; comorbidities: COPH and HF; other comorbidities; CRP; or WBC would improve the accuracy of prediction.




Interaction between MR-proADM and COPD/HF.

## Relationship between Length of Stay and Age

In **Supplementary Figure s1** the relationship between *Length of Stay* and *Age* is shown.

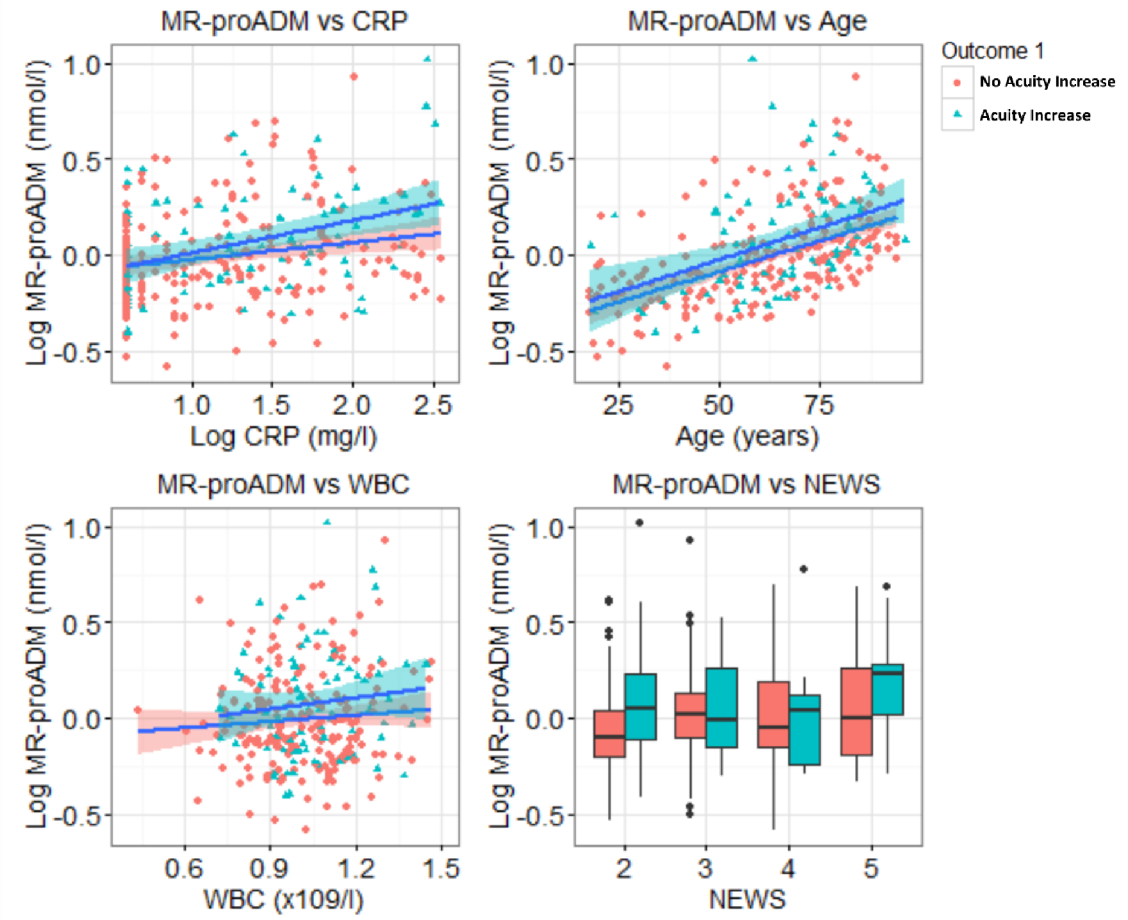



**Supplementary Figure s1**. Relationship between *Length of Stay* and *Age*.

Supplementary information for Graziadio et al, 2018

## Checking for multicollinearity and autocorrelation

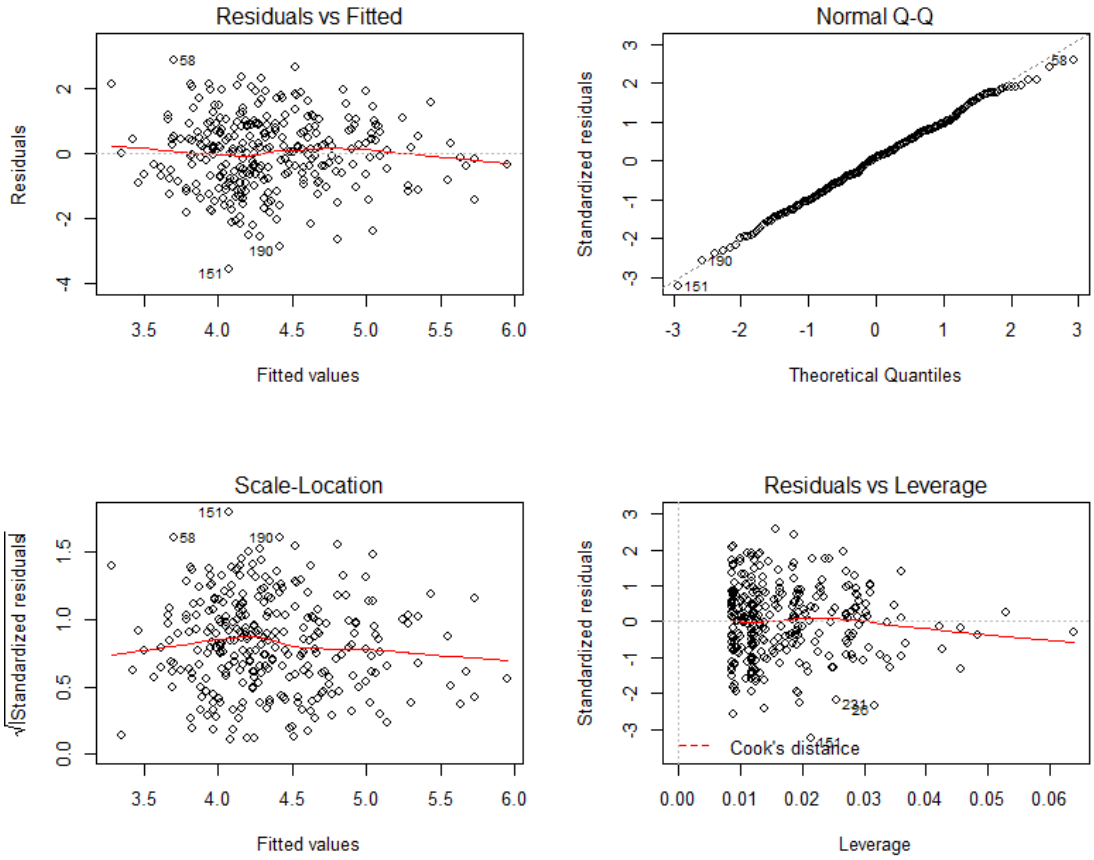
Correlations among biomarkers were also investigated through plotting to evaluate multicollinearity and added value of MR-proADM versus other biomarkers. Plots are shown in **Supplementary Figure s2**.




**Supplementary Figure s2.** Associations between MR-proADM and CRP, age, WBC, and NEWS.

Supplementary information for Graziadio et al, 2018

## Diagnostic plots for length of stay analysis




**Supplementary Figure s3.** Diagnostic plots for linear regression evaluating the prediction accuracy of MR-proADM for *Length of Stay*.

The diagnostics of the model showed no multicollinearity in the data since all the correlation coefficients among the independent variables were smaller than 0.5. No autocorrelation was found in the data, thus residuals are independent from each other: the Durbin-Watson test estimated d = 2.02 (p = 0.56). Evidence for homoscedasticity was provided graphically by the randomly scattered points and almost horizontal fitted lines in **Supplementary Figure s3**, (Residuals vs fitted plot). Analysis of Cook's distance showed that there were no influential points (d <4/51, **Supplementary Figure s3**).

## Analyses of shorter term outcomes

The analyses found that NEWS and MR-proADM had much lower accuracy in predicting *Acuity Increase* at 24 and 12 hours from admission than in predicting *Acuity Increase* at 72 hours as apparent in **Supplementary Table s1**.

Supplementary information for Graziadio et al, 2018

**Supplementary Table s1**. Logistic regression for NEWS and MR-proADM predicting *Acuity Increase* within 24 hours. AIC = 326; AUC = 0.59.

| Covariate | Beta | CI | Odds ratio(CI) | P-value |
|---|---|---|---|---|
| **Intercept** | –1.42 | –1.70, –0.64 | 0.42 (0.27, 0.64) | NA |
| **NEWS 3** | –0.006 | –0.64, 0.71 | 1.04 (0.53, 2.04) | 0.985 |
| **NEWS 4** | –0.37 | –1.58, 0.20 | 0.52 (0.21, 1.22) | 0.368 |
| **NEWS 5** | 0.21 | –0.66, 1.01 | 1.25 (0.52, 2.91) | 0.6244 |
| **MR-proADM** | 0.22 | –0.06, 0.44 | 1.20 (0.95, 1.55) | 0.0547 |

**Supplementary Table s2.** Logistic regression for NEWS and MR-proADM predicting *Acuity Increase* within 12 hours. AIC = 266; AUC = 0.57.

| Covariate | Beta | CI | Odds ratio(CI) | P-value |
|---|---|---|---|---|
| **Intercept** | –1.83 | –2.47, –1.24 | 0.16 (0.08, 0.29) | NA |
| **NEWS 3** | 0.29 | –0.46, 1.04 | 1.34 (0.63, 2.85) | 0.442 |
| **NEWS 4** | –0.15 | –1.57, 0.76 | 0.86 (0.31, 2.14) | 0.756 |
| **NEWS 5** | 0.29 | –0.74, 1.23 | 1.33 (0.48, 3.42) | 0.564 |
| **MR-proADM** | 0.06 | –0.24, 0.31 | 1.06 (0.79, 1.36) | 0.656 |

## Analyses of time-lag effect between news assessment and blood collection for assessment of MR-proADM levels

Given the practicalities involved, it was not possible to stipulate the timings of taking the NEWS on admission and collecting the blood sample for MR-proADM testing. It was expected that difference in times would normally be less than 6 hours, but in 44 subjects the time difference was more than 6 hours.

To investigate the impact of time differences being greater than expected, another analysis was carried out excluding subjects for whom the difference was more than 6 hours (time-lag compliant dataset). The hypothesis was that, if the time difference was an important parameter for the predictive accuracy of MR-proADM level, model coefficients would be greater and confidence intervals narrower for the compliant model. This was not the case; results were similar in the full dataset with 292 subjects and in the compliant dataset with 248 subjects (**Supplementary Table s3**).

Supplementary information for Graziadio et al, 2018

**Supplementary Table s3.** Logistic regression for NEWS and MR-proADM predicting *Acuity Increase* for the time-lag compliant dataset. AIC = 295; AUC = 0.60.

| Covariate | Beta | CI | Odds ratio(CI) | P-value |
|---|---|---|---|---|
| **Intercept** | –1.15 | –1.70, –0.64 | 0.42 (0.27, 0.64) | NA |
| **NEWS 3** | –0.04 | –0.64, 0.71 | 1.04 (0.53, 2.04) | 0.909 |
| **NEWS 4** | –0.65 | –1.58, 0.20 | 0.52 (0.21, 1.22) | 0.152 |
| **NEWS 5** | 0.22 | -0.66, 1.01 | 1.25 (0.52, 2.91) | 0.613 |
| **MR-proADM** | 0.19 | -0.06, 0.44 | 1.20 (0.95, 1.55) | 0.135 |

## Estimation of sensitivities and specificities for each model

For completeness we estimated the sensitivity and specificity for each model in the article (Supplementary Table s4). We used the Youden's index to estimate "optimal" cut-offs. In the next phase of the MR-proADM evaluation, the optimal cut-off will be re-estimated through a decision analysis informed by the role of the test in the pathway.

**Supplementary Table s4.** Sensitivity and specificity of the logistic regression models. Predictor set a. was excluded from the table since the AUROC was too low to calculate meaningful diagnostic accuracy data.

| Models | Sensitivity | Specificity |
|---|---|---|
| ***Acuity Increase*** | | |
| **Predictor set *b*.** | 0.38 (0.29, 0.49) | 0.83 (0.78, 0.88) |
| **Predictor set *c*.** | 0.66 (0.55, 0.76) | 0.69 (0.63, 0.76) |
| ***Deterioration Event*** | | |
| **Predictor set *b*.** | 0.72 (0.56, 0.88) | 0.56 (0.50, 0.62) |
| **Predictor set *c*.** | 0.59 (0.44, 0.75) | 0.83 (0.78, 0.87) |

TRIPOD Checklist: Prediction Model Development

| Section/Topic | Item | Checklist Item | Page |
|---|---|---|---|
| **Title and abstract** | | | |
| Title | 1 | Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted. | P1 L1-3 |
| Abstract | 2 | Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions. | P2 L22-46 |
| **Introduction** | | | |
| Background and objectives | 3a | Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models. | P4-5 L52-84 |
| | 3b | Specify the objectives, including whether the study describes the development or validation of the model or both. | P5 L82-84 |
| **Methods** | | | |
| Source of data | 4a | Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable. | P5 L87 |
| | 4b | Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up. | P5 L87-91 |
| Participants | 5a | Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres. | P5 L87-97 |
| | 5b | Describe eligibility criteria for participants. | P5 L98-102 |
| | 5c | Give details of treatments received, if relevant. | N/A |
| Outcome | 6a | Clearly define the outcome that is predicted by the prediction model, including how and when assessed. | P6-7 L128-149 |
| | 6b | Report any actions to blind assessment of the outcome to be predicted. | P6 L123-126 |
| Predictors | 7a | Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured. | P4 L54-57 P4-5 L70-84 P5-6 L106-117 |
| | 7b | Report any actions to blind assessment of predictors for the outcome and other predictors. | P6 L123-126 |
| Sample size | 8 | Explain how the study size was arrived at. | P5 L92-97 |
| Missing data | 9 | Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method. | P8-9 L197-201 |
| Statistical analysis methods | 10a | Describe how predictors were handled in the analyses. | P7-8 L151-194 |
| | 10b | Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation. | P7-8 L151-194 P7 L153-155 |
| | 10d | Specify all measures used to assess model performance and, if relevant, to compare multiple models. | P7-8 L156-194 |
| Risk groups | 11 | Provide details on how risk groups were created, if done. | N/A |
| **Results** | | | |
| Participants | 13a | Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful. | P10-11 Table 1. |
| | 13b | Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome. | P10-11 Table 1. |
| Model development | 14a | Specify the number of participants and outcome events in each analysis. | P10-11 Table 1. |
| | 14b | If done, report the unadjusted association between each candidate predictor and outcome. | P13-15 Table3&4 |
| Model specification | 15a | Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point). | P13-15 Table3&4 |
| | 15b | Explain how to the use the prediction model. | P15 L244-253 |
| Model performance | 16 | Report performance measures (with CIs) for the prediction model. | P13-15 Table3&4 P15 L244-253 |
| **Discussion** | | | |
| Limitations | 18 | Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data). | P18-19 L306-316 |
| Interpretation | 19b | Give an overall interpretation of the results, considering objectives, limitations, and results from similar studies, and other relevant evidence. | P17-19 L278-323 |
| Implications | 20 | Discuss the potential clinical use of the model and implications for future research. | P19 L325-335 |
| **Other information** | | | |
| Supplementary | 21 | Provide information about the availability of supplementary resources, such as study | See Comments |

| | | | |
|---|---|---|---|
| information | | protocol, Web calculator, and data sets. | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study. | P20 L346 |

Comments:

1. **Item 21: Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.**

Supplementary material - with additional information on methods and results - is attached as separate document. Study protocol and data sets will be available in due course, new project website currently under construction.