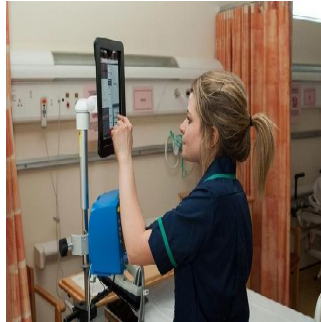


## Assessing Early Warning Score Performance



Sicker patients generally have more vital sign assessments, particularly immediately before an adverse outcome, and especially if the vital sign monitoring schedule is driven by an early warning score (EWS) value. This lack of independence could influence the measured discriminatory performance of an EWS.

According to a new study published in the journal *Resuscitation*, vital sign measurements can be treated as if they are independent – multiple observations can be used from each episode of care – when comparing the performance and ranking of EWSs, provided no EWS includes age.

"Using multiple observations from each episode of care does not significantly change the ranking of EWSs compared to using only one observation from each episode, as long as no EWS includes age. This is in spite of observed dependence between vital signs observations collected during the same episode of care," say researchers from the University of Portsmouth, Portsmouth, UK.

EWS systems allocate points in a weighted manner, based on the derangement of a predetermined set of patient vital signs variables (eg, blood pressure, heart rate, temperature) from an arbitrarily agreed "normal" range. The points for each variable are summed and the total is used to inform a change in the patient's vital sign monitoring schedule and/or trigger a call for expert help at the bedside.

### Methodology

The study analysed data from adult patients (aged 16 years or older) admitted to Portsmouth Hospitals NHS Trust on or after 25/05/2011 and discharged on or before 31/12/2012. The authors used a population of 1,564,143 consecutive vital signs observation sets collected as a routine part of patients' care. Each vital signs measurement set contained: pulse rate, breathing rate, systolic and diastolic blood pressure, temperature,  $S_pO_2$ , the inspired gas (eg, oxygen or air) at the time of  $S_pO_2$  measurement, and the patient's conscious level. Conscious level was recorded as alert (A), responds to voice (V), responds to pain (P) or unresponsive (U).

The authors compared 35 published EWSs for their discrimination of the risk of death within 24 hours of an observation set using (1) all observations in the authors' dataset, (2) one observation per patient care episode, chosen at random, and (3) one observation per patient care episode, chosen as the closest to a randomly selected point in time in each episode.

The authors compared the area under the ROC curve (AUROC) as a measure of discrimination for each of the 35 EWSs under each observation selection method and looked for changes in their rank order.

### Results and Discussion

In the study period, there were 64,285 episodes of care with admission on or after 25/05/2011 and discharge on or before 31/12/2012, where the patient was aged  $\geq 16$ , the patient was not discharged alive on the day of admission and one or more observations were taken during the last 24 hours of the stay. Associated with these episodes of care were 1,395,941 observation sets (mean 21.7 observation sets per episode)

There were no significant changes in rank order of the EWSs based on AUROC between the different observation selection methods, except for one EWS that included age among its components. Whichever method of observation selection was used, the UK National Early Warning Score (NEWS) showed the highest discrimination of risk of death within 24 hours. This supports the findings of earlier studies in which multiple observations were used from each episode of care.

AUROCs were higher when only one observation set was used per episode of care (significantly higher for many EWSs, including NEWS). EWSs that include age, the observation selection method can be important in determining the rank order of EWSs.

### Conclusions

Overall, the findings of this research suggest that vital signs and derived EWS values for EWSs that do not include age can be treated as if they were independent (even though the interclass correlations demonstrate that there is within-episode dependence). Therefore use of multiple observation sets from a single episode in assessing the performance of EWS systems does not appear to bias the ranking of EWSs, as long as no EWS includes age.

Image Credit: UK Engineering and Physical Sciences Research Council

