Universiteit Leiden
The Netherlands

Risk stratification in emergency medicine: towards improved age and sex adjusted risk assessment

Candel, B.G.J.

Citation

Candel, B. G. J. (2024, March 14). *Risk stratification in emergency medicine: towards improved age and sex adjusted risk assessment*. Retrieved from https://hdl.handle.net/1887/3721827

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3721827

Note: To cite this publication please use the final published version (if applicable).

PART IV

CONCLUSIONS


EMERGENCY

10

GENERAL DISCUSSION

GENERAL DISCUSSION

Vital signs serve as the cornerstone of risk assessment of undifferentiated patients presenting to the Emergency Department (ED), but the interpretation of physiological variables deserves more attention. A correct interpretation of physiological variables may lead to earlier treatment (i.e., fluid resuscitation, vasopressors) and better disposition decisions, which may potentially lead to better outcomes, especially in older patients who are at higher risk for adverse outcomes.

In Chapter 1, I have described four objectives of this thesis regarding the age-adjusted interpretation of physiological variables for risk stratification in ED patients, developing a new age- and sex-adjusted risk tool for the hospital, and describing potential bias if risk tools are used for comparing the quality of care among departments. There are several key findings. First, risk tools should consider using age and sex-adjusted numerical scores instead of single cut-off values, as clinically relevant thresholds don't exist for almost all vital signs and biomarkers. Secondly, the International Early Warning Score (IEWS), a recalibrated NEWS including age and sex, substantially improves mortality prediction in the ED with better classification of patients into low and high risk. This score should replace the NEWS for better risk stratification in the ED. Thirdly, standardized mortality ratios (SMRs) are not directly comparable in patients admitted from the ED if hospitals differ in their policy to stabilize patients in the ED before ICU admission. This finding is important if the quality of care is compared among ICU departments.

10.1 Age- and sex-adjusted interpretation of physiological variables

Due to higher baseline risks, absolute mortality risk increases substantially more in older compared to younger patients with changing vital signs and biomarkers [Chapters 2 and 3]. These more marked increases in risk imply that deviating vital signs, such as a low blood pressure, have a different meaning in older patients. For example, a systolic blood pressure (SBP) of 110mmHg has a two times higher odds for mortality than an SBP above 140mmHg. Although these higher odds would still result in a low absolute mortality risk for younger patients, this blood pressure significantly impacts the outcome of elderly patients on cohort level. The same principle applies to other vital signs and most biomarkers; if deviating from their reference values, the impact is more considerable for older patients. Therefore, clinicians should be aware of these differences between young and old patients so that treatment and disposition decisions may be improved in the ED by earlier recognition of vital threads. This awareness can be created using an age-adjusted risk tool such as the International Early Warning Score (IEWS), which we have developed and described in Chapter 8.

Points in numerical risk tools, such as early warning scores, should be based on regression coefficients, which are the logarithm of odds ratios.[175, 204] Because the odds ratios for deviating vital signs for young and old patients are comparable, the same number of points must be assigned for young and old patients for deviating vital signs, as we have done for the IEWS in Chapter 8. For deviating biomarker values, however, young and old patients do not always have similar relative risk increases. For biomarkers of renal function and homeostasis (e.g., urea, creatinine, sodium) and hemoglobin, age should be considered an effect modifier [Chapter 3]. For increasing urea and creatinine, relative mortality has more marked increases in younger than older patients. The same concept is genuine for deviating sodium and hemoglobin levels. Different reference intervals can explain the effect of age on the interpretation of these biomarkers. Both sodium and hemoglobin levels decline with increasing age in a healthy population, whereas reference intervals for urea and creatinine increase due to physiological changes.[62, 76, 77] Remarkably, most laboratories use the same reference intervals for young and old patients. Using age-adjusted reference intervals for these biomarkers may improve the interpretation of test results.[62] More importantly, the effect of age on the associations between these biomarkers and outcomes imply that these biomarkers should be used differently in risk tools for young and old patients. Different points must be assigned depending on age if these biomarkers are used for risk stratification. Because biomarkers may be expensive, not rapidly available or easily repeated, we did not consider them in the development of IEWS.

Male sex was found to be an independent predictor of mortality in different ED settings.[92-95] However, in these studies, it was not reported whether disease severity in males and females was similar at ED presentation. In Chapter 9, we have shown that despite similar characteristics and disease severity at ED arrival, male patients have higher adjusted risks for mortality and high dependency care unit admission than female patients. If healthcare providers are aware of these sex differences in clinical outcomes, they may be able to anticipate the clinical course better and provide better treatment choices for both men and women. It is therefore essential to consider these sex differences in risk stratification. For this reason, we have included sex in the newly developed IEWS.

10.2 Using thresholds for physiological variables

Many acute care guidelines and risk stratification tools use single thresholds for physiological variables, suggesting that one cut-off may discriminate between a good and bad prognosis. However, we have demonstrated in Chapters 2 and 3 that these thresholds do not exist, and mortality increases gradually with deviating vital signs or biomarkers. For mean arterial pressure (MAP), respiratory rate (RR), and temperature, we found single cut-offs at which the prognosis suddenly deteriorates. However, using these single cut-offs would ignore a further increase of risk with more extreme values, so we recommend not using single cut-offs in risk scores. Instead, risk tools should include numerical scores.

Remarkably, an SBP between 110-140mmHg is associated with increased mortality, while an SBP of 120mmHg is considered normal in a healthy population. This result shows that high blood pressure in the ED is normal, probably caused by an adrenergic response due to stress or pain. The absence of this blood pressure response could mean that the patient is critically ill. Another explanation may be that many people have hypertension and thus are used to higher SBP values. As shown in Chapter 4, the difference between the baseline SBP and initial SBP in the ED was associated with 30-day mortality. Therefore, a seemingly normal blood pressure in the ED may in fact be relative hypotension and may not be recognized as such, possibly leading to under-treatment and higher mortality, especially in older patients. Although the study described in Chapter 4 was not designed to assess whether the delta SBP was a better predictor for mortality than the initial SBP in the ED, I believe that the baseline SBP is an important variable that should be considered, if available, as the deviation from a patients' normal SBP may be relevant. In addition, we showed in Chapter 6 that tachycardia may not be helpful to detect hypotension in the ED, which emphasizes the importance to know which blood pressure value is normal for a patient.

10.3 Risk stratification tools

Risk stratification tools are commonly used for risk assessment in the ED. Early warning scores are designed for prognostication and can be used as early as in the ED and add to the clinical impression of a patient's disease severity.[182] However, most early warning scores contain methodological weaknesses in development and validation.[130] Also, they are often not sex- and age-adjusted or they use single thresholds for age in the score. In Chapter 7, we demonstrated that adding age to the National Early Warning Score (NEWS) resulted in better mortality prediction for all age categories. However, in the development of NEWS, points for vital sign disturbances were allocated based on clinical consensus rather than on a statistical approach, which has led to discalibration.[10, 30] For this reason, we have developed the IEWS, a recalibrated sex- and age-adjusted score, which outperforms the most widely adopted NEWS.

Despite their widespread implementation, many professionals are resistant and skeptical towards formal predictions by risk scores, which is a well-known socio-psychological phenomenon.[205] Although one may argue that clinicians could assess patients' risk by experience or clinical gestalt, it is well established that even experts are biased in their perception of risk.[168, 206, 207] In making predictions and judgments under uncertainty, people do not appear to follow the calculus of chance or the theory of statistical prediction. Instead, they rely on a limited number of heuristics which sometimes yield reasonable judgments and sometimes lead to severe and systematic errors.[168, 208, 209] People tend to express great confidence in their biased decisions, called the illusion of validity.[210, 211] For this reason, it is crucial to use a formal risk estimate hand in hand with clinical gestalt. Furthermore, these scores were not designed to replace clinical gestalt but to delay or support it. A low-risk score may allow the clinician more time before assessing the patient, which can especially be beneficial in overcrowded EDs, whereas, for a high-risk score, clinical input may be urgently needed. The IEWS provides an objective risk that may help with clinical decision-making and add to the clinical appearance. The IEWS allows clinicians to select the patients who need an urgent clinical assessment or delay this assessment if the risk of the patient is low. Reducing alarm fatigue is especially important in EDs that become more and more overcrowded. With IEWS, fewer acute assessments are necessary, because many young patients have low risks, whereas older patients are more often at high risk for adverse outcomes. The IEWS, therefore, reduces alarm fatigue compared to NEWS by improving the recognition of patients who really need timely clinical input. This timely input can lead to earlier clinical decisions which may positively affect patients' outcome. Furthermore, it may also help to support the clinician in early initiation of comfort or palliative care if the calculated mortality risk is high in frail elderly.

Using the IEWS may lead to better treatment and disposition decisions. Consequently, this may lead to lower mortality rates and shorter hospital admission periods. Another advantage of a well-calibrated score is that it can be used for case-mix correction when comparing outcomes of patients in national and international research or quality improvement projects, which is already done with the Acute Physiology and Chronic Health Evaluation (APACHE)-IV for ICU patients. However, as we demonstrated in Chapter 9, the calculated severity of illness and predicted mortality risk differ for a given patient if the same initial treatment is given before or after ICU admission or if the ED stay is longer, resulting in a phenomenon called lead-time bias in the literature.[37, 38] These findings imply that a correction or a margin of error should be applied whenever the quality of care among ICUs is compared using the standardized mortality ratio (SMR). Similar lead-time bias might be present in the ED if patients receive different treatments prehospital, depending on the setting. For example, if patients are more often intubated prehospital in one region because of longer distances, the physiological variables measured in the ED will be affected, and so will the predicted mortality risk. With more hospitals joining the Netherlands Emergency department Evaluation Database (NEED), quality of care can be compared among EDs in the near future. Consequently, it is essential to understand differences in the prehospital setting to correct for lead-time bias.

10.4 Limitations

The limitations of each study included in this thesis have been discussed in the relevant chapters. Some general limitations will be mentioned here. Most studies in this thesis used the NEED as the main data source. Due to the retrospective design, we cannot exclude the risk of selection and information bias. For this reason, we have validated many findings in multiple cohorts. For the development of IEWS, we tried to reduce bias by using multiple imputations for missing variables. Furthermore, although we have demonstrated that the IEWS improves mortality prediction in ED patients, the treatment paradox has likely affected our findings due to the retrospective design. This bias rests on the assumption that those at high risk receive aggressive treatment, leading to an apparent overestimation of risks because the outcome is often averted.[30] For example, fever was not associated with higher mortality, as demonstrated in Chapters 2 and 8. A better-provided quality of care, or the treatment paradox, may be partly responsible for this lower risk of patients with fever.[212] However, many other physiological explanations have been suggested that could explain the higher risk for patients with hypothermia, which should not be ignored.[213] Validation of our findings in hospital settings with almost no treatment paradox, such as rural areas in developing countries, will help to understand the treatment paradox’s effect.

10.5 Future perspectives

This thesis mainly focused on investigating differences between men and women and young and old regarding the risk for adverse outcomes. More importantly, we have developed a new tool for prognostication in the ED reducing alarm fatigue and improving prognostication. The next step should be the implementation of the IEWS in Western countries and investigate whether this will lead to better patient outcomes. This implementation may be done together with other measures that might contribute to better recognition of elderly patients at risk. For example, frailty screening instruments, such as the Clinical Frailty Score (CFS) or APOP, may help further guidance of disposition decisions in older patients.[214, 215] However, it is crucial that newly introduced risk tools or screening instruments do not lead to more administration for healthcare workers and should instead reduce workload. I believe that the biggest challenge for the future ED is implementing better risk stratification tools, such as the IEWS, without requiring physicians to calculate the score by multiple clicks or by hand. This could be achieved by implementing the scores automatically in monitors or Electronic Health Records. Prediction models can also be developed using Artificial intelligence (AI). However, so far, AI does not outperform simple logistic regression,[91] and most models are poorly reported and at high risk of bias such that the reported predictive performance is likely optimistic.[216] The advantage of a simple risk score, such as IEWS, is that clinicians understand what the predicted risk is based on. As long as acute care is delivered by humans, not by robots, the latter is essential for humans to adopt the score in clinical practice.

Although we have shown that the IEWS performs well in two Western countries, a prediction model applied in a new healthcare setting or country often produces predictions that are miscalibrated.[164, 190] This may lead to incorrect and potentially harmful clinical decisions.[189] For this reason, further validation is desired to assess the generalizability of the IEWS across multiple settings. In new healthcare settings, the IEWS might need to be updated before it can safely be applied in that new setting.[178] Instead of developing and updating the IEWS in local settings, an international dataset including data from multiple countries and healthcare systems might allow better generalisability and implementation of risk scores, such as IEWS, across different populations.[189] However, this may be a long shot due to privacy and ownership challenges. A more promising solution for better generalisability across multiple settings may be federated learning, a data-private collaborative learning method where multiple EDs could train a machine learning model on their own data at the same time, and then send their model updates to a central server to be aggregated into a consensus model.[217] In this way, data do not need to be shared among institutions. International and interdisciplinary collaborations are necessary for further improvements of prediction models used in the acute care. The NEED could contribute to this fast-growing and developing landscape.