Improving risk prediction model quality in the critically ill: data linkage study

Paloma Ferrando-Vivas,[1] Manu Shankar-Hari,[2,3]
Karen Thomas,[1] James C Doidge,[1]
Fergus J Caskey,[4,5] Lui Forni,[6] Steve Harris,[7,8]
Marlies Ostermann,[9] Ivan Gornik,[10]
Naomi Holman,[11] Nazir Lone,[12] Bob Young,[13]
David Jenkins,[14] Stephen Webb,[15]
Jerry P Nolan,[16,17] Jasmeet Soar,[18]
Kathryn M Rowan[1] and David A Harrison[1*]

[1]Clinical Trials Unit, Intensive Care National Audit & Research Centre, London, UK
[2]Intensive Care Unit, Guy's and St Thomas' NHS Foundation Trust, London, UK
[3]School of Immunology & Microbial Sciences, Kings College London, London, UK
[4]Population Health Sciences, University of Bristol, Bristol, UK
[5]Department of Renal Medicine, North Bristol NHS Trust, Bristol, UK
[6]Department of Clinical and Experimental Medicine, Faculty of Health Sciences, University of Surrey, Guildford, UK
[7]Department of Critical Care, University College London Hospitals NHS Foundation Trust, London, UK
[8]Bloomsbury Institute for Intensive Care Medicine, Division of Medicine, University College London, London, UK
[9]Department of Critical Care, Guy's and St Thomas' NHS Foundation Trust, London, UK
[10]Intensive Care Unit, University Hospital Centre Zagreb, Zagreb, Croatia
[11]Institute of Cardiovascular & Medical Sciences, University of Glasgow, Glasgow, UK
[12]Usher Institute, College of Medicine and Veterinary Medicine, University of Edinburgh, Edinburgh, UK
[13]Diabetes UK, London, UK
[14]Department of Cardiothoracic Surgery, Royal Papworth Hospital NHS Foundation Trust, Cambridge, UK
[15]Department of Anaesthesia and Intensive Care, Royal Papworth Hospital NHS Foundation Trust, Cambridge, UK
[16]Warwick Medical School, University of Warwick, Coventry, UK
[17]Department of Anaesthesia and Intensive Care Medicine, Royal United Hospital Bath NHS Trust, Bath, UK
[18]Critical Care Unit, Southmead Hospital, North Bristol NHS Trust, Bristol, UK

*Corresponding author david.harrison@icnarc.org

Declared competing interests of authors: Manu Shankar-Hari is a member of the National Institute for Health and Care Research (NIHR) Efficacy and Mechanism Evaluation Funding Committee (2020–present) and reports funding through a NIHR Clinician Scientist Award (CS-2016-16-011). Fergus J Caskey reports personal fees from Baxter International (Deerfield, IL, USA) outside the submitted work. Lui Forni reports grants and personal fees from Baxter International and personal fees from Fresenius SE &Co. KGaA (Bad Homburg v.d.H., Germany) outside the submitted work. Naomi Holman reports grants from Diabetes UK (London UK) and NHS England and Improvement outside the submitted work. David Jenkins reports grants from Heart Research UK (Leeds, UK) outside the submitted work. Jasmeet Soar reports personal fees from Elsevier (Amsterdam, the Netherlands) outside the submitted work. Kathryn M Rowan was a member of the NIHR Health and Social Care Delivery Research (formerly Health Services and Delivery Research) Programme Commissioned Board (2014–16) and Funding Committee (2014–19).

Published December 2022
DOI: 10.3310/EQAB4594

Scientific summary

Improving risk prediction model quality in the critically ill

Health and Social Care Delivery Research 2022; Vol. 10: No. 39
DOI: 10.3310/EQAB4594

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

A previous National Institute for Health and Care Research study identified the opportunity to make better use of routinely collected data (both administrative data from death registrations and routine hospital returns and high-quality clinical data from national clinical audits) to better understand the risk factors for, and consequences of, critical illness. Data linkage with routinely collected data sources can provide enhanced information on risk factors and allow exploration of additional outcome measures, leading to improvements in the risk models used to underpin national clinical audits.

Objectives

1. To improve risk models for adult general critical care by (1a) developing risk models for mortality at fixed time points and time-to-event outcomes, developing risk models for longer-term chronic health outcomes of (1b) end-stage renal disease (ESRD) and (1c) type 2 diabetes, and (1d) developing risk models for subsequent health-care utilisation and costs.
2. To improve risk models for cardiothoracic critical care by (2a) enhancing risk factor data and (2b) developing risk models for longer-term mortality.
3. To improve risk models for in-hospital cardiac arrest by (3a) enhancing risk factor data, (3b) developing risk models for longer-term mortality and (3c) developing risk models for subsequent critical care utilisation.
4. Immediate translation of the improved risk models into practice through (4a) adoption into routine comparative outcome reporting for the national clinical audits, and (4b) communication of research output to providers, managers, commissioners, policy-makers and academics in critical care.

Methods and results

Data sources and data linkage

The primary sources of data for this project were the Case Mix Programme (CMP) national clinical audit of adult critical care and the National Cardiac Arrest Audit (NCAA) of in-hospital cardiac arrests. These were linked with data collected for the UK Renal Registry (UKRR), National Diabetes Audit (NDA), National Adult Cardiac Surgery Audit (NACSA), Hospital Episode Statistics (HES) and Office for National Statistics (ONS) death registrations.

The following records were extracted: for CMP, all patients admitted to a participating critical care unit in England between 1 April 2009 and 31 March 2016; for NCAA, all patients experiencing an in-hospital cardiac arrest in a participating hospital in England between 1 April 2011 and 31 March 2016; for UKRR, all patients who started renal replacement therapy (RRT) prior to 31 December 2016 and were alive on 1 April 2009; for NDA, all registrations in audit years 2008–9 to 2015–16; for NACSA, all patients undergoing cardiac surgery between 1 April 2009 and 31 March 2016; for HES, all finished consultant episodes ending between 1 April 2004 and 31 March 2016; for ONS, all deaths registered from 1 April 2009 to 31 March 2016.

NHS Digital, acting as a trusted third party, undertook a bespoke data linkage between the HES/ONS data set and the five national clinical audits. The CMP and NCAA were treated as index data sets. The approvals and data linkage processes were extremely protracted, taking over 4 years from submitting the first data access request to receiving the final linked data set.

Between 1 April 2009 and 31 March 2016, there were 1,007,149 eligible admissions to 248 adult critical care units participating in the CMP. Of these, 965,576 (95.9%) admissions had identifiable links with HES. Between 1 October 2011 and 31 March 2016, there were 89,030 eligible resuscitation team visits following 2222 calls for cardiac arrests reported by 202 hospitals participating in NCAA. Of these, 83,939 (94.3%) had identifiable links with HES.

Mortality after hospital discharge among critically ill patients in England

Patients were included in this analysis if they were discharged alive from hospital between 1 April 2009 and 15 March 2010 following an episode of critical care; the final follow-up date was 15 March 2015. The outcome was time to death following discharge from acute hospital, established by data linkage with death registrations.

Of 50,869 patients discharged alive from hospital, 17,489 (34.4%) died during follow-up. Mortality at 30 days, 90 days, 1 year and 5 years was 2.1%, 4.7%, 11.8% and 32.3%, respectively. Five-year mortality for the age- and sex-matched general population was 10%. Pre-existing risk factors such as age, comorbidities, and functional status had the greatest influence on longer-term outcomes. Acute severity, organ support and length of stay in critical care had comparatively small effects.

Risk models for mortality following admission to adult critical care

Risk prediction models were developed for mortality at 30 days, 90 days and 1 year following admission to critical care, and at 1 year following hospital discharge.

The models for 30-day and 90-day mortality included 119,509 patients admitted to a participating critical care unit between 1 January 2014 and 31 December 2014. The starting point for model development was the previous risk prediction model for acute hospital mortality. All risk factors remained important in predicting 30-day mortality. The final model showed excellent discrimination (c index 0.90) in both internal and external validation. Differences in benchmarking between acute hospital mortality and 30-day mortality were modest and there was little evidence that using a fixed time point reduced heterogeneity. When refitted to 90-day mortality, the relative importance of severe conditions such as metastatic disease and severe liver disease increased.

The model for 1-year mortality following admission included 127,855 patients admitted between 1 January 2013 and 31 December 2013. All risk factors for acute hospital mortality remained important in predicting 1-year mortality. Fewer acute conditions were retained and there were more cancer-related conditions. Most additional comorbidities available via data linkage with HES were important in predicting 1-year mortality; however, the strongest effects remained for the severe conditions already collected in the CMP.

The model for 1-year mortality following hospital discharge included 100,450 patients discharged alive from hospital between 1 January 2013 and 31 December 2013. The effects of comorbidities were largely similar when the model for 1-year mortality was refitted to hospital survivors.

Risk models for development of end-stage renal disease following critical care

Patients were included in this analysis if they were discharged alive from hospital between 1 April 2009 and 31 March 2016 following a critical care episode, excluding those with pre-existing ESRD. The outcome was new receipt of RRT for ESRD following hospital discharge, identified by linkage with UKRR. Death from any cause before ESRD was treated as a competing risk. Cause-specific hazard ratios were estimated using Cox proportional cause-specific regression models, and subdistribution hazard ratios and cumulative incidence functions were estimated using Fine-Gray regression models.

A total of 598,603 patients were included in the analysis. Median follow-up time was 2.7 years and 2831 (0.47%) patients subsequently received RRT for ESRD (1.52 per 1000 person-years follow-up). The strongest predictors were prior hospital admissions involving chronic kidney disease [adjusted

hazard ratio (aHR) 4.11] or acute kidney injury (aHR 1.73) in the preceding 5 years, admission following nephrectomy (aHR 1.92), creatinine during the first 24 hours of critical care (aHR 7.4 for 120 vs. 60 μmol l^{-1}), and duration of renal support (aHR 1.13 per day).

Risk models for development of type 2 diabetes following critical care

Patients were included in this analysis if they were discharged alive from hospital between 1 April 2009 and 31 March 2016 following a critical care episode, excluding those with pre-existing diabetes (type 1 or type 2). The outcome was a new registration for type 2 diabetes, based on the date of diagnosis recorded in the NDA. Death from any cause before a diagnosis of type 2 diabetes was treated as a competing risk. Cause-specific hazard ratios (cHRs), subdistribution hazard ratios (sHRs) and cumulative incidence functions were estimated as for the analysis of ESRD.

A total of 497,967 patients were included in the analysis. Median follow-up time was 2.8 years, and 12,808 (2.6%) patients were subsequently diagnosed with type 2 diabetes (7.8 per 1000 person-years follow-up). The strongest predictors were blood glucose during the first 24 hours of critical care (aHR 3.0 for 12 vs. 8 mmol l^{-1}), pancreatic surgery (aHR 2.83), severe liver disease (aHR 1.60), body mass index (aHR 2.5 for 35 vs. 25 kg/m^2) and Asian (aHR 2.13) and black (aHR 1.43) ethnicities.

Hospital resource use and costs post-critical care

Patients were included in this analysis if they were discharged alive from hospital between 1 April 2013 and 31 December 2014 following a critical care episode. Resource use was measured as the number of days in acute hospital until 1 year following discharge. Total cost was calculated by summing the costs for subsequent hospitalisation and for subsequent critical care admissions, based on health-care resource groups and the Department of Health and Social Care-admitted patient care tariff. A two-part regression model was used to model predictors of costs: a logistic model for any cost versus no cost and a generalised linear model with a gamma distribution and a log link function for the mean cost, conditional on non-zero cost.

A total of 207,805 patients were included in the analysis. The mean health-care cost during the first year after index hospital discharge was £3734. The distribution of total cost was highly skewed with a large mass at zero. A total of 97,593 patients (47%) had a non-zero health-care cost with a mean cost of £7951 (median £4566, interquartile range £2288–9587); 14,293 (6.9%) patients were admitted to critical care during the first year after index hospital discharge, with a mean cost of £9466 (median £4142, interquartile range £2761–9436). Predictors subsequent health-care cost were: previous hospitalisation, critical care length of stay, age, body mass index, illness severity, mechanical ventilation, dependency prior to admission, source of admission, cardiopulmonary resuscitation, deprivation, severe conditions in the past medical history and comorbidities identified by data linkage with HES.

Risk models for adult cardiothoracic care

Patients were included in this analysis if they were admitted to a cardiothoracic critical care unit between 1 April 2009 and 31 March 2015 within 20 days following cardiac surgery, identified by data linkage with NACSA. Risk models were developed for two outcomes: acute hospital mortality and 1-year mortality.

A total of 27,687 patients admitted to seven cardiothoracic critical care units were included in the analysis: 1072 (3.9%) died during the hospitalisation and 1918 (6.9%) died during the 1-year follow-up. The starting point for model development was the previous risk prediction model for acute hospital mortality. In addition to predictors from the previous model, the following factors from NACSA and HES were found to be important: diabetes, atrial fibrillation/flutter, dyspnoea status pre-surgery, history of pulmonary disease, history of neurological dysfunction, extracardiac arteriopathy, operative urgency, cumulative bypass time, severe cardiovascular disease and congestive heart failure. The final model had excellent discrimination (c index 0.89–0.91); however, we found little impact on benchmarking compared with a generic model. Additional predictors in the model for 1-year mortality were renal

function/dialysis, left ventricular ejection fraction, number of previous myocardial infarctions and major aortic procedure.

Risk models for in-hospital cardiac arrest

Prediction models were developed for return of spontaneous circulation (ROSC) sustained for > 20 minutes (ROSC > 20 minutes), hospital survival, 1-year survival and total length of stay in critical care (based on data linkage with CMP).

The risk models for ROSC > 20 minutes, hospital survival and 1-year survival included 26,748 patients experiencing an in-hospital cardiac arrest in one of 172 hospitals between 1 January 2013 and 31 December 2014. The models were validated on 7073 patients experiencing an in-hospital cardiac arrest between 1 January 2015 and 30 June 2015. In the development data set, 12,566 (47.0%) patients achieved ROSC > 20 minutes, 5349 (20.0%) survived to hospital discharge and 4,454 (16.6%) survived to 1 year. The starting point for model development was previous prediction models for ROSC > 20 minutes and hospital survival. All factors from the previous models remained important and the following addition comorbidities, identified from HES, were added: for all three models, congestive cardiac failure, malignancy and metastatic solid tumour; for ROSC > 20 minutes, peripheral vascular disease, diabetes mellitus and chronic renal disease; for hospital survival, peripheral vascular disease, liver disease and hemiplegia or paraplegia; and for 1-year survival, liver disease and chronic renal disease.

A total of 4841 patients were included in the analysis of length of stay in critical care. The mean total critical care unit length of stay was 8 days for hospital survivors and 4 days for non-survivors, with mean costs of approximately £13,000 and £7,000, respectively. For survivors, the following factors were significant in determining total critical care unit length of stay: age, severe conditions in the past medical history, location of arrest, presenting rhythm, reason for admission to critical care by body system, number of advanced organs supports received, Intensive Care National Audit & Research Centre (ICNARC) physiology score, and interactions between severe conditions in the medical history and ICNARC physiology score. For non-survivors, only the following variables significantly influenced the total critical care unit length of stay: age, number of advanced organs supports received; ICNARC physiology score, and interactions between number of advanced organs supports and ICNARC physiology score.

Conclusions

We have successfully linked CMP and NCAA with five other data sources, providing enhancements in risk models for these audits in the form of additional predictors and novel outcome measures. The greatest barriers to maximising the full potential of data linkage were the inordinate amount of time obtaining and maintaining approvals for the use of multiple data sources from multiple data controllers.

Implications for health care

These results have potentially important implications for the future benchmarking of critical care units through the CMP and NCAA. Having demonstrated feasibility of these linkages, ICNARC should investigate cost-effective approaches to routinely link data to support ongoing reporting from the audits. Although comorbidities were found to improve predictions, they had a greater influence on longer-term than shorter-term outcomes. Given the time-lags involved in linking data, we propose that initial quarterly reporting for the audits continue to use directly collected data and that data linkage is undertaken annually to provide enhanced annual reporting including 1-year outcomes.

At the bedside, the new models may assist in providing objective estimates of potential outcomes to patients and their families. A better understanding of factors predictive of worse longer-term outcomes may help to identify those patients requiring greater support in their recovery following critical illness.

Recommendations for research

- Multidisciplinary research should develop and test care pathways for recovery following critical illness using risk models to target those with the greatest need.
- Further relevant data sources for longer-term outcomes following critical illness should be explored, for example stroke.
- Data linkage for resource use and costs following critical illness should be widened to include primary care, outpatient and emergency department data.

Study registration

This study is registered as NCT02454257.

Funding

This project was funded by the National Institute for Health and Care Research (NIHR) Health and Social Care Delivery Research programme and will be published in full in *Health and Social Care Delivery Research*; Vol. 10, No. 39. See the NIHR Journals Library website for further project information.

Health and Social Care Delivery Research

ISSN 2755-0060 (Print)

ISSN 2755-0079 (Online)

Health and Social Care Delivery Research (HSDR) was launched in 2013 and is indexed by Europe PMC, DOAJ, INAHTA, Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and NCBI Bookshelf.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

This journal was previously published as *Health Services and Delivery Research* (Volumes 1–9); ISSN 2050-4349 (print), ISSN 2050-4357 (online)

The full HSDR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hsdr.

Criteria for inclusion in the *Health and Social Care Delivery Research* journal

Reports are published in *Health and Social Care Delivery Research* (HSDR) if (1) they have resulted from work for the HSDR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

HSDR programme

The HSDR programme funds research to produce evidence to impact on the quality, accessibility and organisation of health and social care services. This includes evaluations of how the NHS and social care might improve delivery of services.

For more information about the HSDR programme please visit the website at https://www.nihr.ac.uk/explore-nihr/funding-programmes/health-and-social-care-delivery-research.htm

This report

The research reported in this issue of the journal was funded by the HSDR programme or one of its preceding programmes as project number 14/19/06. The contractual start date was in August 2015. The final report began editorial review in January 2021 and was accepted for publication in March 2022. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HSDR editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health and Care Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the HSDR programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the HSDR programme or the Department of Health and Social Care.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).