



OPEN ACCESS

Evaluation of the DAVROS (Development And Validation of Risk-adjusted Outcomes for Systems of emergency care) risk-adjustment model as a quality indicator for healthcare

Richard Wilson,¹ Steve W Goodacre,¹ Marcin Klingbajl,¹ Anne-Maree Kelly,² Tim Rainer,³ Tim Coats,⁴ Vikki Holloway,⁵ Will Townend,⁶ Steve Crane⁷

¹Medical Care Research Unit, School of Health and Related Research, University of Sheffield, Sheffield, UK

²Western Health and The University of Melbourne, Victoria, Australia

³Accident and Emergency Medicine Academic Unit, Chinese University of Hong Kong, Shatin, New Territories, Hong Kong

⁴Emergency Department, Leicester Royal Infirmary, Leicester, UK

⁵Northampton General Hospital, Northampton, UK

⁶Department of Emergency Medicine, Hull Royal Infirmary, Hull, UK

⁷Department of Accident and Emergency, York District General Hospital, York, UK

Correspondence to

Professor Steve Goodacre, School of Health and Related Research (SCHARR), University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, UK; s.goodacre@sheffield.ac.uk

Received 9 January 2013

Revised 6 February 2013

Accepted 12 February 2013

Published Online First

19 April 2013

ABSTRACT

Background and objective Risk-adjusted mortality rates can be used as a quality indicator if it is assumed that the discrepancy between predicted and actual mortality can be attributed to the quality of healthcare (ie, the model has attributional validity). The Development And Validation of Risk-adjusted Outcomes for Systems of emergency care (DAVROS) model predicts 7-day mortality in emergency medical admissions. We aimed to test this assumption by evaluating the attributional validity of the DAVROS risk-adjustment model.

Methods We selected cases that had the greatest discrepancy between observed mortality and predicted probability of mortality from seven hospitals involved in validation of the DAVROS risk-adjustment model. Reviewers at each hospital assessed hospital records to determine whether the discrepancy between predicted and actual mortality could be explained by the healthcare provided.

Results We received 232/280 (83%) completed review forms relating to 179 unexpected deaths and 53 unexpected survivors. The healthcare system was judged to have potentially contributed to 10/179 (8%) of the unexpected deaths and 26/53 (49%) of the unexpected survivors. Failure of the model to appropriately predict risk was judged to be responsible for 135/179 (75%) of the unexpected deaths and 2/53 (4%) of the unexpected survivors. Some 10/53 (19%) of the unexpected survivors died within a few months of the 7-day period of model prediction.

Conclusions We found little evidence that deaths occurring in patients with a low predicted mortality from risk-adjustment could be attributed to the quality of healthcare provided.

INTRODUCTION

The development of reliable and credible measures to assess the performance of emergency care systems is an important research priority.¹ Mortality is an important outcome in emergency care but differences in crude mortality may reflect differences in case mix rather than quality of care. Statistical models can be used to produce risk-adjusted estimates of mortality that take differences in case mix into account.² These can be used to identify hospitals with high risk-adjusted mortality, with the inference that this reflects suboptimal quality of care.³ This approach was notably used in

the UK to identify potential problems at the Mid Staffordshire NHS Trust.

Risk-adjustment involves measuring patient characteristics that are known to predict mortality and then using these to predict a probability of death for each patient treated by the emergency care system. These probabilities can be summed across a population to estimate the expected death rate for the population overall which can then be compared to the observed death rate. The ratio of the observed to the expected death rate is typically expressed as a standardised mortality ratio (SMR). The UK Department of Health currently uses routinely available data to estimate a Summary Hospital Mortality Index for each hospital in England and Wales.⁴ The DAVROS study (Development And Validation of Risk-adjusted Outcomes for Systems of emergency care) showed that the addition of physiological variables to routine age and diagnostic data could improve prediction and provide excellent discriminant value for 7-day mortality across a range of settings.⁵

The use of risk-adjusted mortality rates as quality indicators assumes that the discrepancy between predicted and observed mortality is attributable to the quality of care provided. If observed mortality exceeds predicted mortality, then this discrepancy is assumed to be due to poor care. Whether this assumption is valid is an important but often neglected aspect of the validation of risk-adjustment methods. Attributional validity refers to that aspect of validating a measure that encompasses testing the assumption that changes seen in the risk-adjusted outcome measure reflect differences in care quality.^{2, 6}

We developed the DAVROS risk-adjustment model to predict mortality in patients admitted to hospital with a medical emergency and generate estimates of risk-adjusted mortality for systems of emergency care. The aim of this study was to evaluate the attributional validity of the DAVROS model in a variety of settings; in other words, to determine whether discrepancies between observed mortality and mortality predicted by the DAVROS model could be explained by the healthcare provided.

METHODS

The methods for derivation and validation of the DAVROS risk-adjustment model have been described in detail elsewhere.⁵ The model was



Open Access
Scan to access more
free content

To cite: Wilson R, Goodacre SW, Klingbajl M, et al. *Emerg Med J* 2014;**31**:471–475.

derived using data from 5644 patients admitted with a medical emergency across three hospitals and then validated using data from 13 762 patients across nine hospitals. The full model used age, International Classification of Disease (ICD-10) code, known malignancy, physiological variables and routine blood data to predict 7-day mortality. A more limited 'physiology' model was developed without the blood variables, which often had high rates of missing data. The c-statistics for the full model ranged from 0.83 to 0.93 across the validation centres, while the c-statistics for the physiology model ranged from 0.80 to 0.91.

In previous studies, attributional validation has used explicit or implicit review.^{2–6} Both involve investigators examining site-level clinical and other material and then reaching a judgement as to the quality of care being delivered. Explicit review involves evaluation of cases against a predefined list of quality criteria to which the investigators refer, so they may literally 'tick the box' if they find the criteria to be present. Implicit review relies on a more general reviewer judgement of quality of care without specific criteria. The process of implicit review is therefore more opaque than explicit review, but allows for more flexibility in an assessment as it is not limited to only predetermined items.

To evaluate the attributional validity of the DAVROS model we used a method that was broadly an implicit review. The review documentation and process allowed our reviewers freedom to interpret and report circumstances as they saw fit with only minimal predetermination of categories, lists or criteria. The study took place in seven hospitals in England, Australia and Hong Kong that participated in the validation phase of the DAVROS study. The research team selected a number of cases (median 30 per hospital) on the basis of having the greatest discrepancy between observed mortality and predicted probability of mortality. Thus, the review-set for each site comprised cases where individuals had died, despite a very low predicted probability of death, and cases where individuals had survived, despite a very low predicted probability of survival.

The reviewer at each site was an experienced emergency physician who worked in the emergency department and acted as investigator during the main data gathering phases of the project. These clinicians reviewed the hospital records and, if appropriate, the death certificate for each case, and identified the factors that they believed significant in explaining the discrepancy between predicted and observed outcome. The clinician reviewers were able to express an opinion as to why the individual outcome was at variance with the original model prediction.

The reviewer completed a form with six questions pertaining to cases of unpredicted death (question set A) and four questions pertaining to cases of unexpected survival (question set B). Question set A or B was completed depending on patient outcome. No personal identifiers were recorded. The questions on the reviewer form are shown in box 1.

Completed forms were mailed back to the research team and data were entered into an Excel spreadsheet for analysis. The responses to questions 2–5 were used to classify deaths as being unexpected due to failure of the model or due to the healthcare system. If question 4 or 5 identified an intervention that potentially contributed to death or an intervention that was not given but could have reasonably been expected to prevent death, then death was classified as being due to failure of the healthcare system. If not, the death was classified as being due to model failure. Descriptive analysis was undertaken to present proportions of cases (with 95% CI) attributable to the healthcare system or aspects of the model.

Box 1 Questions on the reviewer form

The questions on the form were as follows:

Section A (unexpected deaths)

Q1: What was the cause of death (from the death certificate)?

Q2: Was there any potential manifestation of the cause of death in the model variables recorded at presentation?

Q3: Was there any potential manifestation of the cause of death in any presentation variable?

Q4: Did any intervention potentially contribute to death? (Responses: yes; no; unable to say. If yes, text box provided for additional comments)

Q5: Could any intervention have reasonably been expected to prevent death? (Responses: yes; no; unable to say. If yes, text box provided for additional comments)

Q6: Why did the patient die when the model predicted that they would survive? (Free text)

Section B (unexpected survivors)

Q7: What was the reason for admission?

Q8: What was the cause of any abnormal physiological variables in the model?

Q9: Did the patient receive a life-saving intervention? (Responses: yes; no; unable to say. If yes, text box provided for additional comments)

Q10: Why did the patient survive when the model predicted that they would die?

The DAVROS study was approved by the Leeds East Research Ethics Committee and relevant institutional ethics committees in Australia and Hong Kong. The National Information Governance Board (UK) also reviewed and approved the project as it involved using patient identifiable data without consent. However, in the phase of the project reported here, identifiable patient data were only accessed by clinical staff from the relevant hospital.

RESULTS

A total of 280 review forms were dispatched and 232 (83%) were completed and returned. Table 1 shows the number of forms dispatched and returned at each centre, and the SMR for the centre estimated using the DAVROS model. If the SMR is above 100, then observed mortality was higher than predicted. Completion rates for individual centres varied from 68% to 95%. The reason for non-completion was not formally recorded, but informally it appeared that the main reason was failure to locate the hospital notes.

Table 2 summarises the responses relating to unexpected deaths. The response is classified as true if answered 'yes' and false if answered as 'no' or 'unable to say'. The rows show the combinations of responses and the tally gives the total number of cases with each combination. We have excluded combinations with a zero tally.

An intervention was identified that could have contributed to death in 5/179 cases (rows 3 and 5) and an intervention was identified that could have been expected to prevent death in 10/179 cases (rows 1 and 6). Overall, therefore, unexpected deaths were only potentially attributed to the healthcare system in 15/179 cases (8%; 95% CI 5% to 13%).

In 111/179 cases (rows 1 and 2) there was a potential manifestation of the cause of death in the model variables recorded

Table 1 Completion rates for review forms

Centre	SMR	Total N in validation phase	N cases selected	N with forms completed	Died	Survived
A	88.1 (83.9 to 92.3)	2048	95	75 (79)	65	10
B	88.3 (83.8 to 92.8)	1515	37	35 (95)	27	8
C	99.0 (93.9 to 104.0)	1501	30	28 (93)	22	6
D	99.9 (94.2 to 105.5)	1434	34	26 (76)	19	7
E	111.7 (103.8 to 119.5)	1478	32	30 (94)	18	12
F	118.0 (111.5 to 124.5)	1592	31	21 (68)	15	6
G	124.1 (110.3 to 137.9)	1043	21	17 (81)	13	4
Total	–	–	280	232	179	53

SMR, standardised mortality ratio.

at presentation, suggesting that the model variables identified a risk of death but may not have appropriately predicted the risk of death in 62% (95% CI 55% to 69%) of cases. In 129/179 cases (rows 2 and 4) there was a potential manifestation of the cause of death in any variable recorded at presentation, indicating that there was potential for presenting characteristics to predict death in 72% (95% CI 65% to 78%). In 35 cases (20%; 95% CI 14% to 26%) there was no manifestation of the cause of death at presentation and no intervention identified that could have influenced survival, so these deaths appeared to be unpredictable and unexplained.

Table 3 shows the interventions that were identified as having potentially contributed to death or prevented death. The reasons for failing to give an intervention were not usually recorded, and in some cases it may have reflected a deliberate decision to withhold an intervention on grounds of poor anticipated quality of life or respect for patient autonomy.

Free text responses were reviewed for common reasons for failure of the model to predict a high risk of death. This suggested that in patients with pneumonia, stroke or severe or multiple co-morbidities (eg, chronic obstructive pulmonary disease, terminal cancer, renal failure, ischaemic heart disease) the model often failed to predict a high risk of death that was evident on examination of the full case record.

Table 4 compares the attributions between centres. The purpose of this table is not to compare frequencies across the centres (numbers are too small for such an analysis) but to demonstrate that cases implicating the healthcare system are distributed across the centres.

Table 2 Categorisation of responses

	Potential manifestation of the cause of death in a model variable	Potential manifestation of the cause of death in any variable	Intervention could have contributed to death	Intervention could have prevented death	Tally
Row 1	T	T	F	T	6
Row 2	T	T	F	F	105
Row 3	F	T	T	F	2
Row 4	F	T	F	F	24
Row 5	F	F	T	F	3
Row 6	F	F	F	T	4
Row 7	F	F	F	F	35
Total					179

F, false; T, true.

Analysis of the 53 unexpected survivors showed that 26 received a potentially life-saving intervention and 27 did not. The potentially life-saving interventions are shown in table 5.

Further information was provided for 17 of the 27 cases where no potentially life-saving intervention was received to explain why they survived when the model predicted they would die. These are outlined in table 6. In most cases (10/17) the patient died after the 7-day window of model prediction, so the model correctly predicted death but not within 7 days. Survival was attributed to response to treatment in three cases. These may have been more appropriately classified as being attributable to the healthcare system, but we analysed these cases as they were classified by reviewers.

DISCUSSION

If risk-adjusted mortality rates are used to judge the quality of healthcare, then the attributional validity of the risk-adjustment model needs to be demonstrated.^{2–6} This may involve showing

Table 3 Interventions that could have contributed to or prevented death

Possibly contributing to death (Q4)	Possibly preventing death (Q5)
Gastrointestinal bleed following anticoagulant and antiplatelet treatment for myocardial infarction	Thrombolytic therapy not given for massive pulmonary embolism
Deterioration following percutaneous pericardial drainage	No anticoagulant given for pulmonary embolus and no operation for peritonitis
Deterioration associated with high flow oxygen in ambulance	Pacing wire not inserted for heart block
Postoperative bronchopneumonia after operation for fractured neck of femur	No prophylactic anticoagulation given to prevent pulmonary embolism
Delayed diagnosis of perforated duodenal ulcer following steroid therapy for chronic obstructive pulmonary disease	Delayed diagnosis of dissecting thoracic aneurysm
	Emergency laparotomy not performed for peritonitis
	Patient refused antibiotics and blood transfusion
	No prophylactic anticoagulation given to prevent pulmonary embolism
	Anticoagulation for atrial fibrillation not given to prevent stroke
	Thrombolytic therapy not given for myocardial infarction

Table 4 Classification by centre, N (%)

Attribution	Centre							Total
	A	B	C	D	E	F	G	
Healthcare system	3 (5)	2 (8)	4 (18)	2 (11)	2 (12)	0 (0)	2 (16)	15
Model failure	59 (91)	9 (34)	16 (73)	13 (69)	15 (83)	10 (67)	7 (54)	129
Unpredictable death	3 (5)	16 (59)	2 (9)	4 (21)	1 (6)	5 (33)	4 (31)	35
Total	65	27	22	19	18	15	13	179

that discrepancies between model prediction and actual outcome are explained by the healthcare provided. In this respect our analysis failed to demonstrate attributional validity for unexpected deaths. Failure of the model to predict outcome for those who died was often explained by failure of the model variables to appropriately reflect the risk of death and only in a minority of cases could failure of prediction be potentially attributed to the healthcare provided.

Even when unexpected deaths could be attributed to the healthcare provided, this may not have reflected suboptimal care. We did not seek to determine the reasons why potentially life-saving interventions were not used, but it is likely that in a proportion of cases the intervention was deliberately withheld out of respect for patient autonomy or as part of an end of life care plan. There is a growing recognition that inappropriate treatments and interventions are very costly to the health system and confer very little benefit to patients and families. This has resulted in an emphasis on more appropriate end of life care decided in discussion with patients and families. Increasingly, these discussions and decisions are being initiated in emergency departments.

Our findings concur with other studies that have struggled to demonstrate the attributional validity of risk-adjusted mortality for assessing the quality of hospital care.^{7–11} Most previous studies have used explicit review methods to estimate quality of care in terms of explicit standards of care and have evaluated the correlation between quality and risk-adjusted mortality at an institutional level. Our approach of evaluating discrepancies in individual cases has not been widely used.

Evaluation of attributional validity is not often undertaken and methods are poorly developed,^{2, 6} so our inability to demonstrate attributional validity may reflect limitations in our methods rather than limitations of the model or the concept of measuring risk-adjusted outcomes to assess quality. Risk prediction inevitably involves a degree of random error and

Table 5 Potentially life-saving interventions received by unexpected survivors

Intervention	Number of times cited
Intravenous fluids	11
Antibiotics	11
Intensive care, airway or respiratory support	5
Oxygen	3
Anticonvulsants	2
Steroids	2
Diuretic therapy	2
Blood transfusion, bronchodilators, vasodilators, insulin, surgical procedure, temporary pacing wire	1 × each

imprecision. Failure of the model to predict outcome may simply reflect this imprecision, particularly in the cases where we identified no manifestation of the cause of death in any presenting variable or any potential healthcare factor. Another limitation is inherent to the review process, which can only capture gross evidence of suboptimal healthcare. More subtle evidence of suboptimal care, such as poor hygiene, poor nursing care or inadequate monitoring, would not be apparent from case note review. We were unable to review a proportion of cases due to inability to retrieve case notes. This may be a source of bias as there may be a specific reason why case notes are not available. For example, if there is an enquiry into the healthcare provided, notes may not be available.

A specific limitation of our methods was that the reviewers were clinicians responsible for care in the hospitals they were reviewing. When asked to review cases from their own institution, clinicians may interpret outcomes in the most favourable or positive light. This may explain why healthcare interventions were more frequently identified as explanations for unexpected survival than unexpected death, and may mean that we have underestimated the proportion of unexplained deaths that could be attributed to healthcare intervention. It could also be argued that the reviewer's judgements may have been influenced by their involvement in developing the DAVROS model study.

Table 6 Explanations identified for unexpected survival

Response to treatment	Responded well to antibiotics and medical treatment Successful medical treatment Early access to consultant surgeon, well monitored, intravenous fluids, antibiotics and insulin early on
Died after 7 days	Responded to medical treatment but then died 2 weeks later Care of dying pathway initiated, discharged to nursing home where he died 2 weeks later Died on day 11 Died 12 days after admission Died 8 days later Died 9 days later In hospital for 6 weeks and then died 4 months after admission In hospital for 4 months. Subsequently died 2 months after discharge Patient died 2 months later Responded to medical treatment. Subsequently died 3 months after admission
Unclear	Unclear. Limited therapy instituted consisting of oxygen, antibiotics and fluids Unclear. Limited treatment given with agreement of family
Poor model prediction	Model prediction of death based on relatively minor derangements of blood variables Variables in the model altered by his particular presentation (infective exacerbation of chronic obstructive pulmonary disease) and hence the likelihood of death 'overstated'

Employing external reviewers, not connected with the institution, would have provided a greater degree of impartiality. However, data protection legislation presents barriers to those from outside the care team accessing notes. Obtaining consent or anonymisation of notes is unlikely to be feasible. Even if data protection barriers can be overcome, the process of independent external review may be prohibitively expensive. Alternatively we could have used an explicit review process in which our reviewer evaluated cases against a predefined list of quality criteria. This could have reduced the potential for reviewers to make biased judgements but would have limited their ability to identify issues that were not predefined. We chose to use an implicit review process because the potential range of judgements that would need to be predefined for an explicit review process would have been substantial and unwieldy, and yet might still not have covered key issues.

Although failure to demonstrate attributional validity undermines the use of risk-adjusted mortality as a quality indicator, it does not mean that risk-adjusted data are worthless. This study identified a number of cases where healthcare was potentially suboptimal and could form the basis for more detailed investigation. Risk-adjustment can provide a means of identifying cases with the greatest discrepancy between predicted and actual outcome, where detailed audit would be most worthwhile. This use of risk-adjustment could be enhanced by further research. For example, this study identified presentation variables that are not currently included in the model, such as ECG abnormalities, that could improve model prediction. In conclusion, however, this study found little evidence that deaths occurring in patients with a low predicted mortality from risk-adjustment could be attributed to the quality of healthcare provided.

Acknowledgements We thank Jon Nicholl for help with developing the idea for the study and the DAVROS Research Team for their help with the project. The DAVROS Research Team includes the Project Management Group (SWG, RW, Neil Shephard, Jon Nicholl, Martina Santarelli, Jim Wardrope); the principal investigators (Alison Walker (Yorkshire Ambulance Service), Anne Spaight (East Midlands Ambulance Service), Julian Humphrey (Barnsley District General Hospital), Simon McCormick (Rotherham District General Hospital), A-MK (Western Hospital, Footscray, Victoria), TR (Chinese University of Hong Kong), TC (Leicester Royal Infirmary), VH (Northampton General Hospital), WT (Hull Royal Infirmary), SC (York District General Hospital)); and the Steering Committee (Fiona Lecky, Mark Gilthorpe, Enid Hirst, Rosemary Harper).

Contributors SWG designed the study. RW managed the study and oversaw data collection. MK, A-MK, TR, TC, VH, WT and SC were responsible for data collection. RW and SWG analysed the data. RW wrote the first draft of the paper. All authors assisted in the interpretation of data and revising the paper, and approved the final draft.

Funding The DAVROS project was funded by the Medical Research Council. The researchers were independent from the funders. The funders had no role in conducting the study, writing the paper, or the decision to submit the paper for publication.

Competing interests None.

Ethics approval Leeds East Research Ethics Committee.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>

REFERENCES

- 1 Department of Health. *Revision to the operating framework for the NHS in England 2010/11*. London: Department of Health, 2010.
- 2 Iezzoni LI. Validity and reliability of risk adjusters. In: Iezzoni LI, ed. *Risk adjustment for measuring health care outcomes*. 3rd edn. Chicago: Health Administration Press, 2003:207–21.
- 3 Jarman B, Gault S, Alves B, et al. Explaining differences in English hospital death rates using routinely collected data. *Br Med J* 1999;318:1515–20.
- 4 Campbell MJ, Jacques RM, Fotheringham J, et al. Developing a summary hospital mortality index: retrospective analysis in English hospitals over five years. *BMJ* 2012;344:e1001.
- 5 Goodacre S, Wilson R, Shephard N, et al. Derivation and validation of a risk-adjustment model for predicting seven-day mortality in emergency medical admissions. *BMJ* 2012;344:e2904.
- 6 Silva LK. Validity of the risk adjustment approach to compare outcomes. *Cad Saude Publica* 2003;19:287–95.
- 7 Thomas JW, Holloway JJ, Guire KE. Validating risk-adjusted mortality as an indicator for quality of care. *Inquiry* 1993;30:6–22.
- 8 Hartz AJ, Gottlieb MS, Kuhn EM, et al. The relationship between adjusted hospital mortality and the results of peer review. *Health Serv Res* 1993;27:765–77.
- 9 Dubois RW, Rogers WH, Moxley JH, et al. Hospital inpatient mortality. Is it a predictor of quality? *N Engl J Med* 1987;317:1674–80.
- 10 Park RE, Brook RH, Koseoff J, et al. Explaining variations in hospital death rates. Randomness, severity of illness, quality of care. *JAMA* 1990;264:484–90.
- 11 Keeler E, Rubenstein L, Kahn KL, et al. Hospital characteristics and quality of care. *JAMA* 1992;268:1709–14.