Journal of Biomedical Informatics 46 (2013) 837–848

Contents lists available at SciVerse ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Development and validation of a continuous measure of patient condition using the Electronic Medical Record

Michael J. Rothman [a,*], Steven I. Rothman [a,b], Joseph Beals IV [a]

[a] PeraHealth, Inc., 1520 S. Boulevard, Suite 228, Charlotte, NC 28203, USA
[b] F.A.R. Institute, Inc., 5019 Kestral Park Dr., Sarasota, FL 34231, USA

ARTICLE INFO

Article history:
Received 11 January 2013
Accepted 25 June 2013
Available online 3 July 2013

Keywords:
Health status indicators
Deterioration
Nursing assessments
Electronic health records
Acuity score
Patient condition

ABSTRACT

Patient condition is a key element in communication between clinicians. However, there is no generally accepted definition of patient condition that is independent of diagnosis and that spans acuity levels. We report the development and validation of a continuous measure of general patient condition that is independent of diagnosis, and that can be used for medical-surgical as well as critical care patients.

A survey of Electronic Medical Record data identified common, frequently collected non-static candidate variables as the basis for a general, continuously updated patient condition score. We used a new methodology to estimate in-hospital risk associated with each of these variables. A risk function for each candidate input was computed by comparing the final pre-discharge measurements with 1-year post-discharge mortality. Step-wise logistic regression of the variables against 1-year mortality was used to determine the importance of each variable. The final set of selected variables consisted of 26 clinical measurements from four categories: nursing assessments, vital signs, laboratory results and cardiac rhythms. We then constructed a heuristic model quantifying patient condition (overall risk) by summing the single-variable risks. The model's validity was assessed against outcomes from 170,000 medical-surgical and critical care patients, using data from three US hospitals.

Outcome validation across hospitals yields an area under the receiver operating characteristic curve (AUC) of $\geqslant$0.92 when separating hospice/deceased from all other discharge categories, an AUC of $\geqslant$0.93 when predicting 24-h mortality and an AUC of 0.62 when predicting 30-day readmissions. Correspondence with outcomes reflective of patient condition across the acuity spectrum indicates utility in both medical-surgical units and critical care units. The model output, which we call the Rothman Index, may provide clinicians with a longitudinal view of patient condition to help address known challenges in caregiver communication, continuity of care, and earlier detection of acuity trends.

1. Introduction

The increasing adoption of Electronic Medical Records (EMRs) enables the consistent and accurate manipulation of large amounts of clinical data for computational purposes. This provides an opportunity to use a wide range of clinical variables to quantitatively determine the acuity of a hospitalized patient's condition without risking the burden of complexity or miscalculation on the part of clinicians [1–4].

Clinicians can currently choose among many tools [4] designed to quantify some aspect of patient acuity. In almost all cases these tools assess risk for a specific event such as cardiopulmonary arrest, mortality, or transfer to the ICU [5–11], specific disease or procedure [12–16] or are specific to an environment, such as PRISM [15], APACHE III [17] and SOFA [18] in the ICU.

These systems were developed using one of two approaches. In some cases researchers compute acuity based on a set of easy-to-apply rules that are derived from expert opinion; the Modified Early Warning Score (MEWS) is a well-known example of this approach [19]. In other cases researchers have developed models using standard regression methods applied against specific events, such as unexpected death or transfer to the ICU [20,21].

1.1. Expert-based models for triggering transfer to a higher level of care

One characteristic of MEWS [19] and related scores such as VitalPAC [22], PEWS [2], Bedside-PEWS [23], C-CHEWS [24], is that they depend on experts to specify the input variables and the risk

* Corresponding author.
E-mail addresses: rothman@aol.com, mjrothman@earthlink.net, mrothman@perahealth.com (M.J. Rothman).

1532-0464
http://dx.doi.org/10.1016/j.jbi.2013.06.011

associated with particular values or states of each input variable. For example, the Behavior/Neuro sub-score used in the C-CHEWS model is assigned a value of zero when the patient state is "playing/sleeping appropriately", a value of 1 when "sleepy, somnolent when not disturbed", a value of 2 when "irritable, difficult to console", and a value of 3 when displaying "reduced response to pain". There are about 40 values in the model that have similarly assigned risk weightings. These expert distinctions form the basis of this model. In general their implied risk functions do not have any rigorous validation, although some authors follow iterative processes based on the model's performance [24]. Difficulties arise in application of expert-based models; for example, to identify 44% of transfers to ICU occurring in the next 12 h, MEWS generates 69 false positives for every correctly identified event [19].

1.2. Regression models for predicting death or an unexpected transfer to higher level of care

A more rigorous approach is to build a regression model, which is then evaluated on an independent test set. Kirkland et al. [20] have taken a parsimonious modeling approach (using only 4 parameters) to predict within 12 h: Rapid Response Team calls, unplanned transfers to the ICU, or unexpected death (AUC = 0.71). Escobar et al. [21] have taken advantage of a large dataset from 14 hospitals to develop a model using approximately 100 data items, from which are derived about 40 regression parameters. The three outcomes which they chose to regress against were, within the next 12 h: unplanned transfer from ward to ICU; or unplanned transfer from transitional care unit (TCU) to the ICU; or sudden unexpected death on ward or TCU without a "do not resuscitate" order in place (AUC ranged from 0.57 to 0.84 depending on the disease specific sub-model). A problem with both these regression models is the use of infrequent heterogeneous events as targets, which lessens their practical applicability and lowers their signal to noise ratio. For example, Escobar makes a direct comparison with the MEWS statistic above: to identify 44% of transfers to ICU occurring in the next 12 h, his model generates 34 false positives for every correctly identified event.

1.3. A Patient condition model for earlier detection of deteriorating trends

Rather than attempting to forecast a particular adverse event, we argue that intervention during early deterioration can help prevent such an adverse event from occurring [25,26]. Quantifying each patient's condition on a continuous basis could provide an opportunity to detect a declining trend in time for clinicians to act, enabling an appropriately graded response to changes in patient condition [26,27].

We report the development and validation of a general measure of individual patient condition using 26 clinical variables commonly available in the EMR including vital signs, lab results, cardiac rhythms and nursing assessments. We describe our approach to variable selection, the evaluation of risk represented by the value of each variable, and the creation of a heuristic model to compute overall risk. We have named the resulting score of patient condition the "Rothman Index" in memory of Florence A. Rothman, whose death inspired this work.

Leveraging the EMR realizes the vision [4,28] of a continuously updated patient condition score independent of specific events, diseases, procedures or environments, and incorporating sufficient clinical variables to provide sensitivity to patient risk across the spectrum of acuity, from the unimpaired to the gravely ill. In developing this patient condition model we have taken a new methodological approach to quantify patient risk by following a data-intensive process that avoids the ad-hoc nature of MEWS-type models while using more frequent outcome events to assign risk values to the model inputs.

2. Material and methods

This section presents the methodology to create and validate a general risk model. We describe criteria for variable selection, the approach to determining risk associated with each candidate variable, and how the selected inputs are combined into a final model. As we are seeking to model "patient condition", a quantity that does not have a defined reference standard, we validate the reported model by showing correlation with quantities that are related to the patient's general condition [29].

2.1. Data

We used a single dataset for model construction and five datasets for model validation. Data access was in some cases opportunistic, leading to variation in dataset sizes and time frames.

Model construction used data for 22,265 in-patients admitted from January 2004 through December 2004 at Sarasota Memorial Hospital (SMH), an 806-bed regional medical center.

Model validation used in-patient data from 3 hospitals: 19,055 admissions from July 2005 to June 2006, 32,341 admissions from September 2007 through June 2009, and 45,771 admissions from January 2008 through May 2010 at SMH; 32,416 admissions from July 2009 to June 2010 at Abington Memorial Hospital (AMH), a 665-bed regional referral center and teaching hospital; and 19,402 admissions from July 2008 to November 2008 at Hospital C, an 800-bed teaching hospital in the southeastern US. In two instances (discharge disposition and APACHE III correlation), we used the SMH 2004 model construction dataset for validation.

Data at all three hospitals were recorded in the same type of EMR system, Allscripts (Eclipsys) Sunrise Clinical Manager. Data from pediatric (<18 years of age), obstetric, psychiatric, were excluded. All other in-patients were included. One-year post-discharge mortality was determined using the Social Security Administration's Death Master File.

This work received independent approval from the Institutional Review Board at each hospital.

2.2. Variable selection and excess risk conferred

A survey of the EMR data collected from SMH yielded approximately 7000 variables (nearly 6500 flowsheet inputs and 500 laboratory tests). However, the aim of a general, continuous, and sensitive measure of patient condition imposed certain constraints upon the variables, that they be: (a) related to patient condition, (b) regularly collected on all patients, and (c) susceptible to change over the course of a patient's hospital stay. Hence, demographic or descriptive variables that do not change during a patient's time in the hospital, such as age, sex, diagnosis, comorbidities, and hospitalization history were excluded. The focus is on "how the patient is" rather than "who the patient is". These constraints reduced the dataset to 43 candidate variables: 13 nursing assessments, 6 vital signs, 23 laboratory tests, and cardiac-monitoring rhythms.

For each of these 43 clinical variables, an excess risk function was computed. "Excess risk" is defined as the percent increase in 1-year all-cause mortality associated with any value of a clinical variable, relative to the minimum 1-year mortality identified for that variable. For example, Fig. 1 shows the excess risk function for white blood cell count (WBC). The points represent average excess 1-year post-discharge all-cause mortality vs. average WBC at discharge; data from 22,265 patient discharges from SMH (2004) are bucketed by WBC range. The regression line is a polynomial

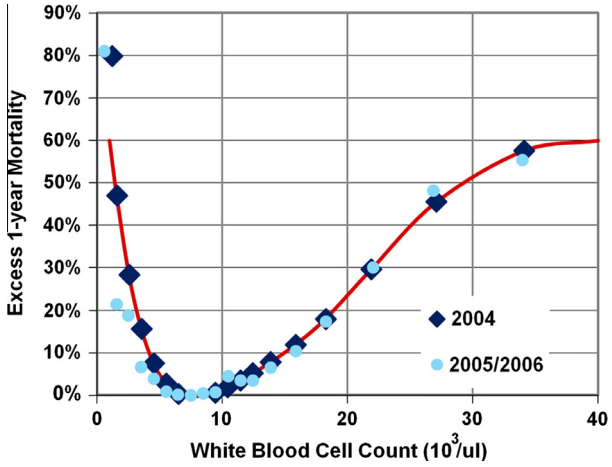

Fig. 1. Excess 1-year mortality risk as a function of white blood cell count. For values of 1 or below, the value is set at 80%. For values of 34 and above, the value is set at 60%. Diamonds show 2004 data, circles show 2005–2006 data.

fit to the data, normalized to the lowest risk value. Above and below clinical value extrema, where data are sparse, the function is set to a constant. Although risk almost certainly increases at values beyond the data extrema, we choose not to extrapolate owing to the fact that the actual slope of the risk function in these domains is unknown. As excess risk functions are important building blocks of our model, in Fig. 1 we also show data points from a second 1-year period, to show the consistency of the underlying relation.

The goal is not to predict mortality; rather, 1-year post-discharge mortality is chosen as an outcome that is well defined, easy to determine, encompasses a significant part of the sample population (approximately 10%), and is related to in-hospital patient condition at the time of discharge [30,31]. Excess risk functions were calculated for the numerical, continuous clinical variables including vital signs and laboratory measures.

Excess risk values were also determined for nursing assessment results. Nursing assessment data are collected in the course of a "head-to-toe" or "body system" patient examination [32] performed at least once per nursing shift and recorded in the EMR in one of two ways. If charting by exception, [33] the nurse answers a master question for each physiological system, such as "*Is the patient's respiratory function within normal limits?*" ("normal" limits might be defined as respiration at 12–24 breaths/minute, nail beds pink, bilateral breath sounds). Alternatively, the nurse may answer a series of questions, such as, "*What are the breath sounds?*", and "*What color are the nail beds?*". The answer to a master question is "pass" or "fail", and when there are multiple questions per assessment, we map the entire category to a "fail" if any answer reflects a deviation from normal. Assessment questions may vary between hospitals but share the aim of noting non-normal physiological system fundamentals. Example definitions of standards for each nursing assessment are shown in Table 1.

Excess risk for each nursing assessment category was calculated from the difference in 1-year mortality between patients who passed and patients who failed their last assessment prior to discharge. In Fig. 2, we show excess risk computed for 12 nursing assessments (excluding the Braden score) for 2 separate 1-year periods at SMH to show the stability of the relative impact of failing a particular nursing assessment. The 2004 results were used in the model development.

With all 43 variables on a common 1-year excess mortality risk scale, multi-collinearity was determined using Pearson correlation coefficients. If any pair of variables had a Pearson correlation coefficient greater than 0.7, the less frequently collected variable was excluded. The relative importance of the remaining variables was determined using forward stepwise logistic regression (SAS Version 9.2) of the excess risk values against 1-year mortality. Variables were added subject to the added regression coefficient having a p-value of less than 0.05. The final set of variables is shown in Table 2. The logistic regression is used only to select variables; its coefficients are not used. The resultant model is hence not a regression model trained on 1-year mortality but rather a heuristic model built using excess risk functions. See Appendix A for the risk functions associated with the selected variables.

2.3. Model construction

The variable selection and the transformation of raw variables into risk functions are conducted using empirical methods; however, the model itself is partially heuristic. As there is no general quantitative measure of patient condition, regression methods are not applicable.

The score is indexed from 100 and reduced as a function of increasing risk. Risk is calculated as the sum of the excess risk represented by each individual variable at a given time, as shown in following equation:

$$\text{Rothman Index} = 100 - (\text{Scale factor}) \sum_{\text{Input}=1}^{\#\text{Variables}} \text{Excess risk}_{\text{Input}} \quad (1)$$

A score of 100 is achieved only when all input variables are at a minimum (zero excess risk) value. A scaling factor ensures the majority of patients on a general medical-surgical unit fall within a dynamic range from 0 to 100, rendering subtle deterioration easily detectable as a falling RI score. Critically ill patients may have negative RI values (the minimum possible RI score is −91). As it is rare that all 26 variables are measured at the same time, the model must allow for missing data. We address this by using the most recent value of each variable when computing the RI, limiting

Table 1
Nursing assessment standards.

Cardiac standard	Pulse regular, rate 60–100 BPM, skin warm and dry. Blood pressure less than 140/90 and no symptoms of hypotension.
Food/nutrition standard	No difficulty with chewing, swallowing or manual dexterity. Patient consuming >50% of daily diet ordered as observed or stated.
Gastrointestinal standard	Abdomen soft and non-tender. Bowel sounds present. No nausea or vomiting. Continent. Bowel pattern normal as observed or stated.
Genitourinary standard	Voids without difficulty. Continent. Urine clear, yellow to amber as observed or stated. Urinary catheter patent if present.
Musculoskeletal standard	Independently able to move all extremities and perform functional activities as observed or stated (includes assistive devices).
Pain standard	Without pain or VAS (visual analogue pain scale) <4 or experiencing chronic pain that is managed effectively.
Neurological standard	Alert, oriented to person, place, time, and situation. Speech is coherent.
Peripheral/vascular standard	Extremities are normal or pink and warm. Peripheral pulses palpable. Capillary refill <3 s. No edema, numbness or tingling.
Psychosocial standard	Behavior appropriate to situation. Expressed concerns and fears being addressed. Adequate support system.
Respiratory standard	Resp. 12–24/min at rest, quiet and regular. Bilateral breath sounds clear. Nail beds and mucous membranes pink. Sputum clear, if present.
Safety/fall risk standard	Safety/fall risk factors not present. Patient is not a risk to self or others.
Skin/tissue standard	Skin clean, dry and intact with no reddened areas. Patient is alert, cooperative and able to reposition self independently. Braden scale >15.

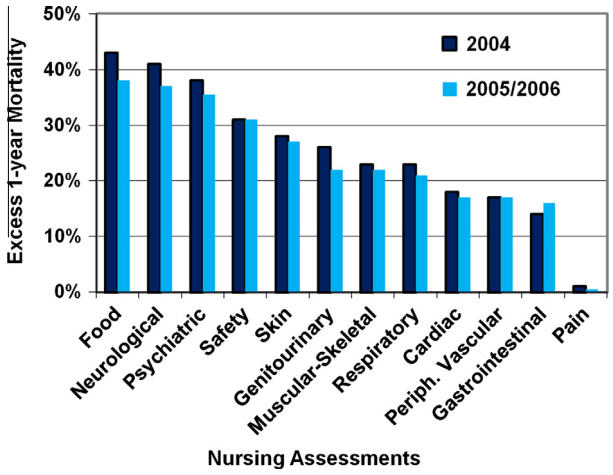

Fig. 2. Excess 1-year mortality risk for each of 12 simplified nursing assessments. The graph shows both data from 2004 and data from July 2005–June 2006 at SMH.

the acceptable time that a measurement can be carried forward (e.g. to 15 h in order to span nursing shifts). If a variable is completely missing for a particular patient, zero excess risk is assigned.

Laboratory tests are generally collected less frequently than vital signs and nursing assessments. To take advantage of the information from laboratory tests without sacrificing accuracy over time, the RI model is composed of 2 sub-models (RI_{noLab} and $RI_{withLab}$). Both sub-models are computed as in Eq. (1); RI_{noLab} uses only nursing assessments and vital signs; $RI_{withLab}$ uses nursing assessments, vital signs and laboratory tests. As the laboratory data ages its relevance to the patient's current condition diminishes; therefore, $RI_{withLab}$ is blended by a linear decay with RI_{noLab}; after 48 h, RI_{noLab} is used solely, until new laboratory data becomes available. At a minimum, computing a patient's RI requires a set of vital signs and nursing assessments.

Stepwise forward logistic regression is used to select variables for each sub-model. Variables selected for RI_{noLab} include all inputs listed in Table 2, except for the laboratory results. The same procedure for $RI_{withLab}$ yielded 24 variables; with the inclusion of lab data, the cardiac rhythms and peripheral vascular inputs do not meet the minimum criterion ($p < 0.05$) for inclusion in the model.

The model is thus a simple linear combination of the two sub-models as a function of time, based on the most recent available laboratory data, as shown in following equation:

$$\text{Rothman Index} = \left[RI_{\text{noLab}}\left(\frac{\text{TimeSinceLabs}}{48}\right)\right] + \text{Smoothing function}\left[RI_{\text{withLab}}\left(1-\frac{\text{TimeSinceLabs}}{48}\right)\right] \quad (2)$$

where "TimeSinceLabs" has a maximum value = 48 h.

This approach allows the lab results to smoothly and gradually "age-out" as they became too far removed in time to be relevant. When new lab data arrive again, then the $RI_{withLab}$ sub-model is selected. A "Smoothing Function", shown in Eq. (2), and described in Appendix B, was added to enhance continuity when RI switches from RI_{noLab} to $RI_{withLab}$.

2.4. Model assessment and construct validators

The Rothman Index is a heuristic model and as such is not designed to predict a specific quantity. Because neither validity nor reliability can be exactly quantified [29,34], we use the method of construct validity as used by Richardson et al. in the development of the Score for Neonatal Acute Physiology. [29] Construct validity is specifically defined by Boudreaux et al. [35] as "... the degree to which a measure actually assesses the attribute it is purported to measure" based on "whether the measures relate to other variables in expected and predictable ways". In this work the relationship of the RI to an outcome independently associated with patient condition was examined on the assumption that poorer condition is expected to correspond to poorer outcomes.

Construct validators were chosen to assess patient condition over different time frames and acuity levels. Sensitivity and specificity are important discriminatory parameters [3,36], and area under the receiver operating characteristic curves (AUC) for several construct validators are reported on a by hospital basis.

2.4.1. Discharge disposition

Discharge disposition provides a method for ascertaining how consistently the RI corresponds to patient condition across the acuity spectrum of discharged patients. The average final RI prior to discharge for patients was compared to each of 6 categories of discharge disposition: home, home healthcare, rehabilitation (outpatient facility or rehabilitation ward), skilled nursing, hospice, and death. Analysis of variance (ANOVA) procedures were employed to determine if statistically significant differences existed between the RI averages associated with discharge categories. The ANOVA was followed with pair-wise contrasts using the Tukey–Kramer HSD test. [37]. Additionally, the final RI was used to separate the first 4 categories from the last 2 categories (patients dying or discharged to hospice) for all three validation hospitals, and the associated AUC was calculated. Patients who left the hospital against medical advice or were transferred to another hospital (e.g. a psychiatric facility) were excluded.

2.4.2. 24-h Mortality

The RI score is calculated every time a new model input is available. Many RI scores are thus recorded for each patient's hospital stay. When the patient is in critical care, data are collected more frequently. Every RI score for every patient was categorized as being within 24 h of death or not; this measures the effectiveness

Table 2
Twenty-six variables chosen as inputs to the RI.

Vital signs	Nursing assessments (head-to-toe)	Nursing assessments (other)	Laboratory tests (blood)	Cardiac rhythm
Temperature	Cardiac	Braden score	Creatinine	Asystole
Diastolic blood pressure	Respiratory		Sodium	Sinus rhythm
Systolic blood pressure	Gastrointestinal		Chloride	Sinus bradycardia
Pulse oximetry	Genitourinary		Potassium	Sinus tachycardia
Respiration rate	Neurological		BUN	Atrial fibrillation
Heart rate	Skin		WBC	Atrial flutter
	Safety		Hemoglobin	Heart block
	Peripheral Vascular			Junctional rhythm
	Food/Nutrition			Paced
	Psychosocial			Ventricular fibrillation
	Musculoskeletal			Ventricular tachycardia

of the RI in discerning patients who are critically ill. AUC for this 24-h mortality correlation was calculated for each hospital.

2.4.3. 30-day Readmissions

The relationship between final RI prior to discharge and 30-day hospital readmission rate was investigated by computing AUC. To explore the sensitivity of the RI to variations in condition for relatively well (low acuity) patients, only readmission rates for patients discharged to home or home healthcare were used in the calculation of the AUC.

2.4.4. APACHE III initial estimate of ICU mortality

To examine high-acuity patients, the RI just prior to admission to ICU was compared to APACHE III initial estimates of mortality [17] upon entry to the ICU, using the SMH 2004 dataset. In this dataset there were 804 patients with APACHE III scores, for whom there was also enough data to compute RI scores. Patients entering from operating, recovery or emergency rooms were excluded because data required to compute RI were unavailable.

2.4.5. Modified Early Warning Score (MEWS)

RI scores and MEWS values were computed for 32,416 AMH patients (yielding more than 1.86 million scores) to contrast the sensitivity of the two scores across the acuity spectrum. MEWS calculations were based on the model of Subbe et al. [19]. A mapping was done from the Glasgow Coma Score to approximate the Alert/Painful/Verbal/Unresponsive (AVPU) input, which was not available in the dataset [38].

3. Results

The results show the correspondence of the RI to patient outcomes across a range of acuity levels. Validation using outcomes across hospitals yields an area under the receiver operating characteristic curve (AUC) of $\geqslant 0.92$ when separating selected patient discharge categories, and an AUC of $\geqslant 0.93$ when predicting 24-h mortality, and an AUC of 0.62 when predicting 30-day readmissions. To compare with an accepted ICU scoring model, we correlate our results with APACHE III estimates of mortality. To compare with an expert-based system for hospital wards, we calculate MEWS scores and show that the RI provides a fine gradation of patient condition.

3.1. Discharge disposition

Table 3 shows the mean, standard deviation, and percentage of patients by discharge category for SMH, AMH, and Hospital C. The mean values of the final RI show a decreasing rank order corresponding to discharge categories associated with increasing patient acuity. The mean RI values for each discharge category are similar across the 3 demographically varied institutions.

All pair-wise comparisons of means were statistically significant ($p < 0.05$) using the Tukey–Kramer HSD test. When the final RI was used to separate the first 4 categories in Table 3 from the last 2 categories (hospice and death), the AUC was 0.923 (95% CI, 0.915–0.930) at SMH and 0.965 (95% CI, 0.960–0.970) at AMH and 0.915 (95% CI, 0.900–0.931) at Hospital C. A statistically significant difference was found between RI values associated with the discharge disposition categories $F(5, 22{,}265) = 4665.0$, $p < 0.0001$, at SMH, see Fig. 3.

A small number of patients in Fig. 3 (3.9% of total deaths) died with RI scores higher than 65. Chart reviews revealed that these relate to sudden cardiac or pulmonary death. Patients discharged to home with a low RI score typically represent instances when care was continued outside the hospital, e.g. outpatient treatment for congestive heart failure patients. Similar results were obtained at all hospitals: $F(4, 32{,}104) = 8204.23$, $p < 0.0001$ for AMH and $F(5, 18{,}809) = 1643.41$ $p < 0.0001$ for Hospital C.

Table 3
Mean Final Rothman Index by Discharge Disposition. Mean final Rothman Index values are in bold, with standard deviation in parentheses, followed by the percent of the population in that discharge category. Abington Memorial Hospital does not have a discharge designation for "home healthcare".

Discharge category	Sarasota Memorial $N = 22{,}265$	Abington Memorial $N = 32{,}104$	Hospital C $N = 18{,}809$
Home	**79.2** (14.0) 74.3%	**82.1** (11.4) 76.1%	**85.5** (10.7) 83.7%
Home healthcare	**70.3** (17.6) 2.8%	Unavailable	**82.0** (11.2) 2.9%
Rehabilitation	**63.6** (15.4) 5.3%	**68.9** (15.7) 2.6%	**74.1** (13.9) 1.5%
Skilled nursing	**60.7** (17.8) 11.3%	**63.1** (16.0) 17.0%	**66.8** (16.0) 10.0%
Hospice	**35.1** (24.1) 4.1%	**35.0** (21.0) 2.4%	**49.2** (23.0) 0.9%
Death	**15.3** (23.6) 2.1%	**13.5** (20.9) 1.9%	**22.9** (27.3) 1.1%

3.2. 24-h Mortality

The average likelihood of death within 24 h as a function of the RI is shown in Fig. 4a. The AUCs relating the RI to 24-h mortality are: 0.933 (95% CI, 0.931–0.936) for AMH; 0.948 (95% CI, 0.945–0.951) for SMH; 0.929 (95% CI, 0.919–0.940) for Hospital C. The figure shows a non-linear relationship between a falling RI and the percentage of patients who die. A 10-point RI decrease when the patient is at 90 has little impact on risk, while a 10-point RI reduction from 20 to 10 reflects a large increase in risk. This fits a general understanding that patients in poor condition have a reduced physiological reserve; [18] hence a disproportionately increased risk is associated with additional impairment.

3.3. 30-Day readmissions

Fig. 4b relates the final RI at discharge to the likelihood of hospital readmission within 30 days for patients discharged to home

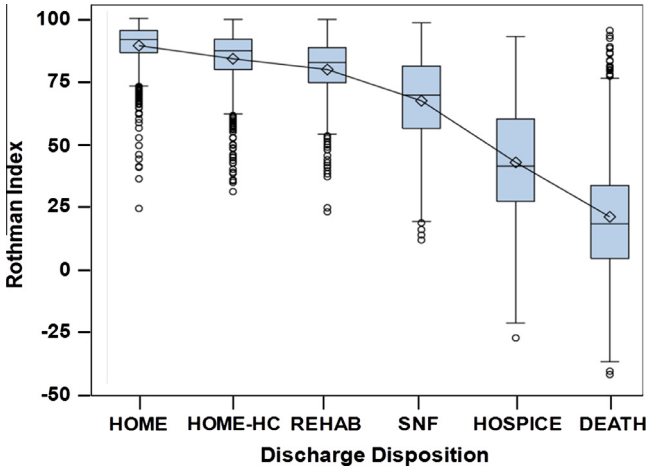

Fig. 3. Final RI vs. Discharge Disposition. Distribution of patients' final Rothman Index prior to discharge vs discharge categories: Home, Home Healthcare, Rehabilitation, Skilled Nursing Facility, Hospice, and Death. The plot is constructed as follows: the top and bottom of each box are at the 75% and 25% percentiles of RI for that discharge distribution. Horizontal lines within each box are median RI values; diamonds show mean RI values. The top of each whisker represents the maximum value or the median plus 1.5 times the interquartile distance (75th percentile–25th percentile); the bottom represents the minimum value or the median minus 1.5 times the interquartile distance. Circles are outliers of the distributions.

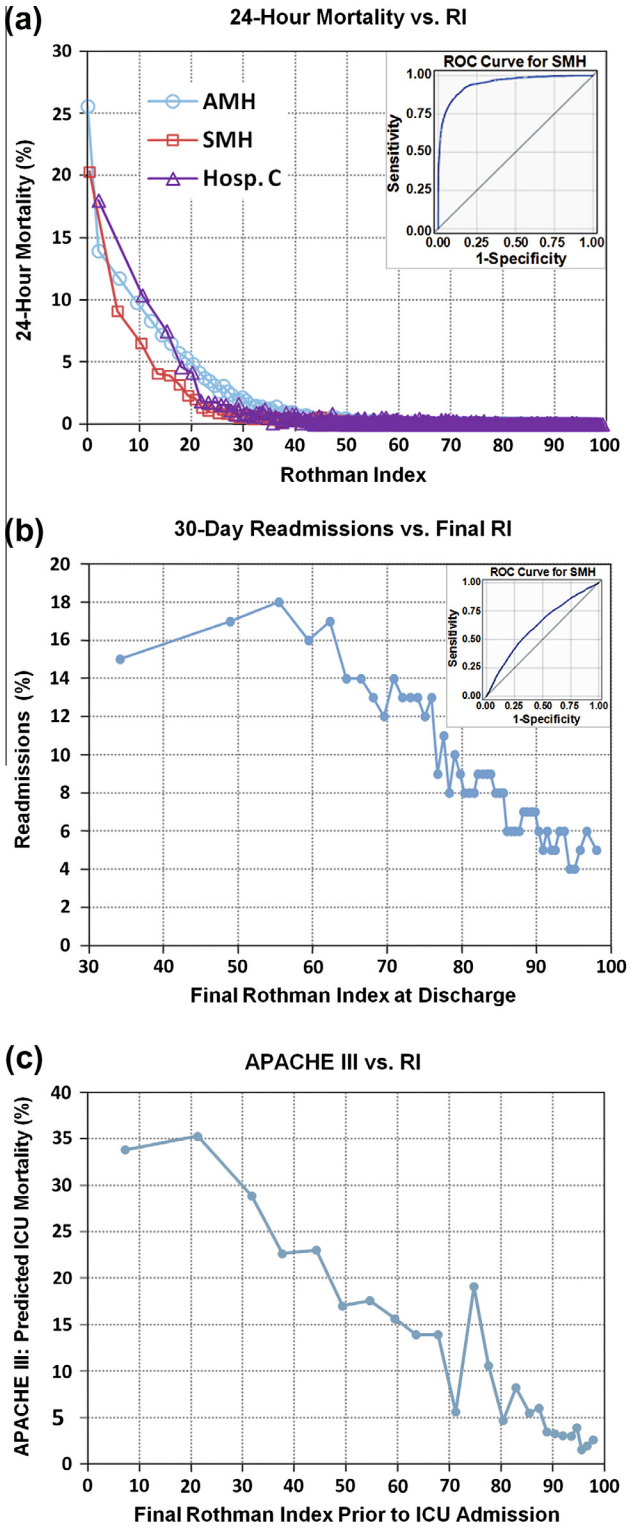

Fig. 4. (a) 24-h Mortality as a function of the Rothman Index; Abington Memorial Hospital (32,416 admissions from 7/2009 to 6/2010, blue circles), Sarasota Memorial Hospital (32,341 admissions from 10/2007 to 6/2009, red squares), and Hospital C (19,402 admissions from 7/2008 to 12/2008, purple triangles). Inset is the corresponding ROC curve for SMH. (b) Percent readmissions within 30 days for patients discharged to home or home healthcare vs the final Rothman Index prior to discharge. The data are from 45,771 admissions to SMH between 1/2008 and 5/2010. Data are plotted in buckets each representing approximately 2% of the population. Inset is the corresponding ROC curve for SMH (c) APACHE III initial estimate of mortality vs final Rothman Index measured prior to entering the ICU. Data for 804 patients are plotted in buckets of approximately 32 patients, from the SMH 2004 dataset. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

or home healthcare. The fall-off in readmissions for RI values below 30 results from the increased mortality rate of these patients. These numbers underestimate the actual 30-day readmission rate as only patients returning to SMH could be identified, thus omitting any readmission events at other hospitals. The AUC was 0.62 (95% CI, 0.61–0.63).

3.4. APACHE III estimate of mortality

The final RI prior to transfer into the ICU for 804 SMH patients from the SMH 2004 dataset was compared with the corresponding APACHE III mortality estimates as shown in Fig. 4c. A significant correlation was found (Pearson's correlation = −0.47, $p < 0.001$). While we do not replicate APACHE III, which has different diagnosis-specific inputs and a goal of predicting death in the ICU, we do demonstrate that low RI values correspond to high APACHE III initial estimates of mortality in the ICU and high RI values correspond to low APACHE III initial estimates of mortality.

3.5. Modified early warning score

The MEWS system is designed to highlight critical deteriorations that may precede cardiac or pulmonary arrest; however, more subtle changes that continually occur in the majority of the patient population are largely invisible to MEWS. RI scores and MEWS values computed on a common data set show that 93% of MEWS readings (Fig. 5) are 0, 1, 2 or 3; these values, which typically fall below the clinical response threshold, are seen in shades of blue) [19,39].

4. Discussion

We have developed a continuously updated index using a novel methodology, based upon 1-year mortality. Asynchronous vital signs, cardiac rhythms, lab tests and nursing assessments are taken as inputs, and multiple validation tests across several independent test sets at three US hospitals are reported. The RI surpasses MEWS in predicting 24-h mortality, correlates with APACHE III in the ICU, and also is sensitive across the full acuity spectrum, with an average final value prior to discharge properly ordered by acuity of discharge dispositions.

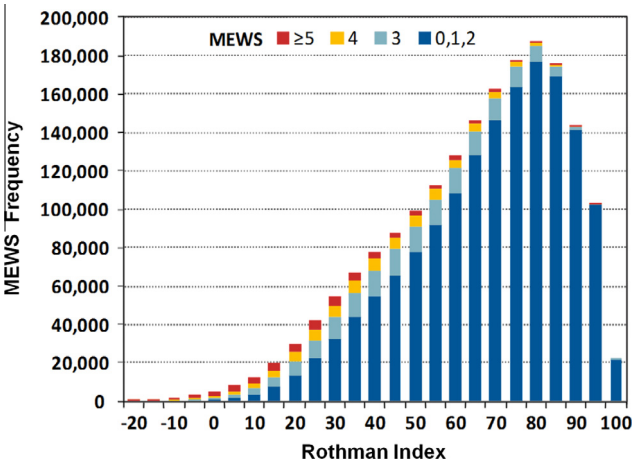

Fig. 5. MEWS frequency vs. RI for 1.86 million scores computed during 32,416 patient visits at AMH. 62% of computed MEWS are "0" or "1", 22% are "2", and 9% are "3".

4.1. Patient condition – meaning and consistency

There is no generally accepted definition of patient condition that is independent of diagnosis and which spans acuity levels. Nevertheless, patient condition is a crucial concept referenced by healthcare providers and is a key element in the communication between clinicians. The approach taken here follows a methodology which recognizes that, while patient condition is difficult to quantify, death serves as a well-defined state that is the farthest point from good condition. This contrasts with approaches that rely on other outcomes, such as transfer to the ICU or RRT calls, neither of which indicates a well-defined state. [21] We have computed excess risk, based upon a comprehensive set of clinical variables, in an effort to quantify distance from death. If a patient has no excess risk, he or she is as far from death as possible at that time and hence in good condition.

Instead of in-hospital death, which is relatively rare (approximately 1–2% of patients), our model is based on 1-year post-discharge mortality, where death is far more common (approximately 10% of patients). This provides five to ten times the signal strength with which to determine the relationships between clinical measures and risk. Possibly other tests could have been substituted. The requirement is that the test be consistently applied across all variables and that the outcome is both sufficiently frequent and a plausible surrogate for condition.

As a generalized assessment of a hospitalized patient's condition, the RI is not designed to predict any specific event. However, we have tested correspondence with immediate measures such as discharge disposition, and future events such as 30-day readmissions and 24-h mortality. We additionally test the RI through comparison with existing validated scores that provide a proxy for current patient condition (e.g. APACHE and MEWS).

The reported construct validation demonstrates a clear correspondence between the RI and events indicative of patient condition, independent of variations in diagnosis, demographics, and acuity, and also conforms to clinical expectations associated with different degrees of acuity, thus substantiating the content and face validity of the RI. [40] The discharge disposition and 24-h mortality validation is based on an independent SMH data set, as well as AMH and Hospital C data sets, representing 3 independent test sets.

Our purpose of using 30-day readmissions for only those patients discharged to home and home healthcare is to complement the other validations: we wanted a measure of low acuity situations. None of our other validations focus on the large majority of patients who have recovered from their illness, and are going home from the hospital. Our AUC of 0.62 is comparable with models designed exclusively to predict 30-day readmission [41–43].

The consistency of the relationships between RI and discharge disposition and RI and 24-h mortality across hospitals is notable, given varying approaches to documenting nursing assessments, and different patient demographics and providers. Furthermore, while the three hospitals in this work have the same type of EMR, our use of common clinical data allows the RI model to be employed regardless of how the information is recorded in any particular EMR system. We believe this establishes the generality of the RI as a transportable measure of patient condition. We have also implemented the model at hospitals using four other EMR systems.

A particular application of a continuous score of current patient condition lies in assisting physicians or nurses in assessing a patient, as well as in considering any assessment as part of a series of assessments to discern a trend. Progression of the patient's general condition over time can be plotted. Hence a graphical depiction of a patient's Rothman Index may provide valuable context for clinicians seeking to understand and communicate the general state of a patient (see Appendix C for examples).

4.2. Input considerations

Variations in data collection frequency and lag between the time a clinician takes a measurement and enters it into the EMR leads to asynchronous data input. [44] The RI accommodates this by recalculating when any input variable changes, regardless of whether other inputs are modified. The RI is more robust due to partial redundancy: for example, respiration information is captured by a nursing assessment and also by a vital sign. These input variables are not collinear and are not interchangeable, they contribute different information about the patient's respiratory state and may be measured at different times. Vital signs are taken throughout the day, usually at 4-h intervals. Nursing assessments are determined once per shift, usually every 12 h. Since we are trying to provide a longitudinal tracking view of the patient's condition, we need to incorporate data as soon as they become available.

One of the inputs to the RI calculation is pulse oximetry (blood oxygen saturation). The RI model does not distinguish between patients with and without supplemental oxygen. The model simply reflects the patient's current condition without regard to what that condition might be if circumstances or treatments were otherwise. This approach eliminates uncertainty or complications that can arise, as with the Worthing score, [45] when oxygen saturation must be ascertained with the patient on room air, potentially inconveniencing the provider and adding risk for the patient [3].

The inclusion of selected laboratory values when available adds sensitivity to the score, while the 48-h time-decay function imposed on the lab values accounts for their diminishing relevance and reliability over time. However, it is recognized that not all lab values have identical durations of relevance, and a more sophisticated treatment of lab value inputs, which allows for different decay-functions or phase-out periods for each of the lab components, may enhance the contribution of lab values to the model.

4.3. Contribution of nursing assessments

Patient deterioration is not always apparent to a physician or nurse, particularly when there is no prior familiarity with the patient. Sensitivity to deterioration, especially when not evidenced by compromised vital signs, is thus crucial for effective patient monitoring in a general medical-surgical unit. Nursing assessments that span the body's physiological systems have high responsivity to a wide range of modes of patient deterioration (see Table 1). Similar to observations of functional measures [31,46], these data can improve sensitivity to early deterioration [47–49] and are particularly relevant to non-critical patients in a general medical-surgical unit. The RI is the first score to incorporate nursing assessments to assist in identifying patient deterioration by capturing "softer fails", such as a patient who is not eating, or is confused (see Table 1). Nursing assessments have been shown to be an important source of dynamic clinical information. The functional deteriorations they reflect may precede vital sign changes and have been found to correlate strongly with in-hospital mortality [30].

The impact of each category of inputs in computing the index can be estimated from the maximum RI point loss possible from each category as a percentage of the total model possible point loss. The contributions of each variable category in RI_{noLab} are: nursing assessments (47%), cardiac rhythms (11%) and vitals (42%). For $RI_{withLab}$ they are: nursing assessments (34%), vitals (35%), and lab results (31%). These numbers can be derived from the excess risk functions shown in Appendix A.

4.4. Sensitivity relative to APACHE and MEWS

Scores designed to predict specific critical events focus on the relevant set of variables providing the greatest discriminatory power in these instances; [5] however, this focused selection of inputs limits such scores' sensitivity to patient risk in other domains of acuity. This is illustrated by APACHE III, which predicts the likelihood of in-ICU mortality. [17] The APACHE III algorithm does not assign additional mortality risk to patients whose temperature is between 96.8°F and 103.8°F. Although this may be appropriate when assessing ICU mortality risk, temperatures within that range are of interest when caring for patients in a general medical-surgical unit. Similarly, MEWS is developed to identify patients at risk for cardiac or pulmonary arrest, with little capacity to capture more subtle changes in condition.

Numerous aggregate weighted track and trigger systems and MEWS variants have been validated against 24-h mortality, but have shown a mixed ability to discriminate between survivors and non-survivors. [50] Previous efforts to take advantage of EMR data to improve the sensitivity and specificity of MEWS systems have fallen short. [51] In contrast, this work has found excellent 24-h mortality AUC values for all 3 hospitals ($\geqslant$0.929). This compares favorably with the recently reported VitalPAC Early Warning System (ViEWS) AUC of 0.888 (95% CI, 0.880–0.895), [22] though it is noted that differences in RI and ViEWS AUC values may be in part due to the scores being calculated with different data sets.

4.5. Limitations

The RI is unavailable for obstetric, psychiatric, and pediatric patients (<18 years of age) because of lack of data at SMH. Investigation into the relationship between risk and clinical values in the pediatric population is currently underway at several children's hospitals.

Data sets obtained from each of the three hospitals do not reflect patient visits over identical periods of time. Additionally, validation analyses for 30-day readmission rates and APACHE III and MEWS score correlations were each restricted to populations from a single hospital (SMH, SMH, and AMH respectively) owing to the limited availability of 30-day readmission, APACHE, and MEWS data at the other hospitals.

Availability and accuracy of social security numbers may limit the determination of 1-year mortality, which is used in the excess risk functions. As a test, we note that for patients with numbers who were discharged to hospice, 98% matched as having died within 1 year of discharge; since we do not expect everyone in hospice to have died, this is an indication that our 1-year mortality estimates are good.

While the RI is intended for automated real-time computation, it can be subject to a time lag owing to the fact that data are not always entered into the EMR promptly upon measurement.

Further, while we have found nursing assessment data valuable, they are subject to a certain subjective variability. The extent of variation and the degree to which nurse education impacts consistency is a topic of ongoing investigation.

Clinical studies using the RI need to be completed in order to ascertain the impact of this tool, in practice, on improving communication and quality of care.

5. Conclusions

The RI is a general measure of patient condition spanning the acuity spectrum, from relatively well patients, to acutely ill patients. It is computed using data commonly available in hospital EMRs, specifically 26 clinical metrics including vital signs, laboratory results, cardiac rhythms and particularly nursing assessments, which are an important source of dynamic clinical information. It is independent of diagnosis, demographics, or comorbidities. Leveraging data now electronically available in the EMR enables the RI to be updated continually throughout the duration of a patient's hospital stay.

The RI has been tested using construct validators and demonstrates excellent correlation with measures reflective of patient condition. Validation work was conducted using data from 170,000 in-patient admissions at 3 US hospitals. This technology may provide clinicians with a longitudinal view of patient condition to help address known challenges in caregiver communication and the early detection of deterioration.

6. Statements for journal publication

6.1. Conflict of interest

The authors are employees and shareholders of PeraHealth, Inc. of Charlotte NC, a company that uses the Rothman Index within its products, which are currently installed at several major medical research centers and hospital systems. The modeling methodology described herein was performed prior to the formation of the company. PeraHealth did not manage, approve or affect any aspect of the writing of this article. Research contributing to improvements in the Rothman Index is being conducted both by the company and by an independent nonprofit foundation, the F.A.R. Institute of Sarasota, FL.

6.2. Access to the Rothman index

The authors will provide access to software for computation the Rothman Index, at no cost to qualified researchers.

Acknowledgments

The authors express their deep appreciation to the staffs of Sarasota Memorial Hospital, Abington Memorial Hospital, and Hospital C. Particular thanks are extended to G. Duncan Finlay, MD, without whose support this work would not have been possible, as well as to George Almasi, PhD, Daniel Rothman, MBA, Alan Solinger, PhD, and Steve Lascher, PhD, for their expert input and advice. Research was funded by grants from the Sarasota Memorial Healthcare Foundation and the Goldsmith Fund of the Greenfield Foundation.

Appendix A. Excess risk functions

Excess risk functions are based on 1-year all-cause mortality as a function of pre-discharge values of the measurements specified. The minimum 1-year mortality is subtracted from all points in order to show incremental risk above baseline.

Each excess risk function reported was constructed using the approach described in Methods section. In a few instances, the function fitted to the data points falls slightly below zero, in which case the risk at those values is set to zero. As the data become sparse about the extrema of the measurements, the risk is set to a constant; we do not extrapolate beyond the actual data. The piecewise-continuous functions which constitute this set of excess risk functions (Fig. A1) are only used to interpolate.

Appendix B. Derivation of the smoothing function

Our approach is based upon the assumption that there is a single number that characterizes a quantity called "patient condition". We use two sub-models to allow us to use laboratory data that, while frequently collected, are not collected as frequently as vital signs. Each sub-model, independent of the number of variables it contains, should measure this same patient condition. The model that uses lab data should be slightly more sensitive, as it incorporates more indicators for assessing patient condition.

To estimate patient condition we compute overall risk, computed by summing the risk inherent in each individual variable. Since the two sub-models have different numbers of variables (RI_{noLab} has 19 and $RI_{withLab}$ has 24) combining them requires a scaling factor so that the overall risk computed by each is on a common scale.

As we described above, RI gradually and smoothly shifts from the $RI_{withLab}$ model to the RI_{noLab} model as a function of "time since the lab data have arrived", over 48 h. However, if the latest lab data are more than 48 h old, there is an abrupt shift to $RI_{withLab}$ when new lab data arrive. Although both models are measuring the same quantity, when switching between sub-models, information may in fact be present in $RI_{withLab}$ which was not contained in RI_{noLab}, causing an appropriate jump or drop in RI. However, it is also possible that in switching from RI_{noLab} to $RI_{withLab}$ there may be a small jump or drop in RI not indicative of a change in patient condition, but that is simply an artifact of combining the two sub-models. To match the sub-models, a smoothing function is applied, which is the variable scaling factor shown in Eq. (2).

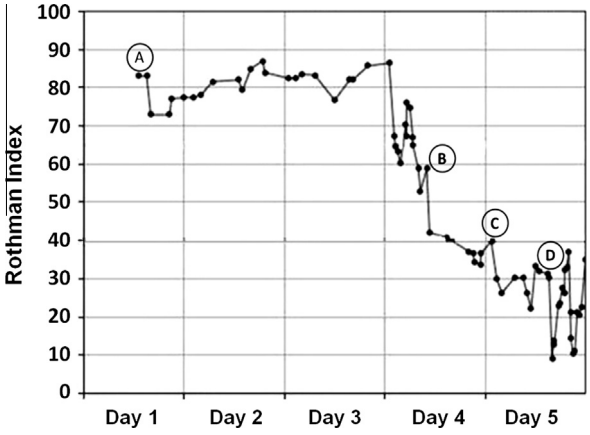

Fig. C1. Rothman Index graph. Each point is a computed RI score. Each vertical line is at midnight. (A) Patient admitted for laparoscopic surgical procedure. (B) Nurse notes patient deterioration; majority of nursing assessments are failed. (C) Rapid Response Team called; patient treated and remains on floor (D) Rapid Response Team called again; patient transferred to ICU with sepsis.

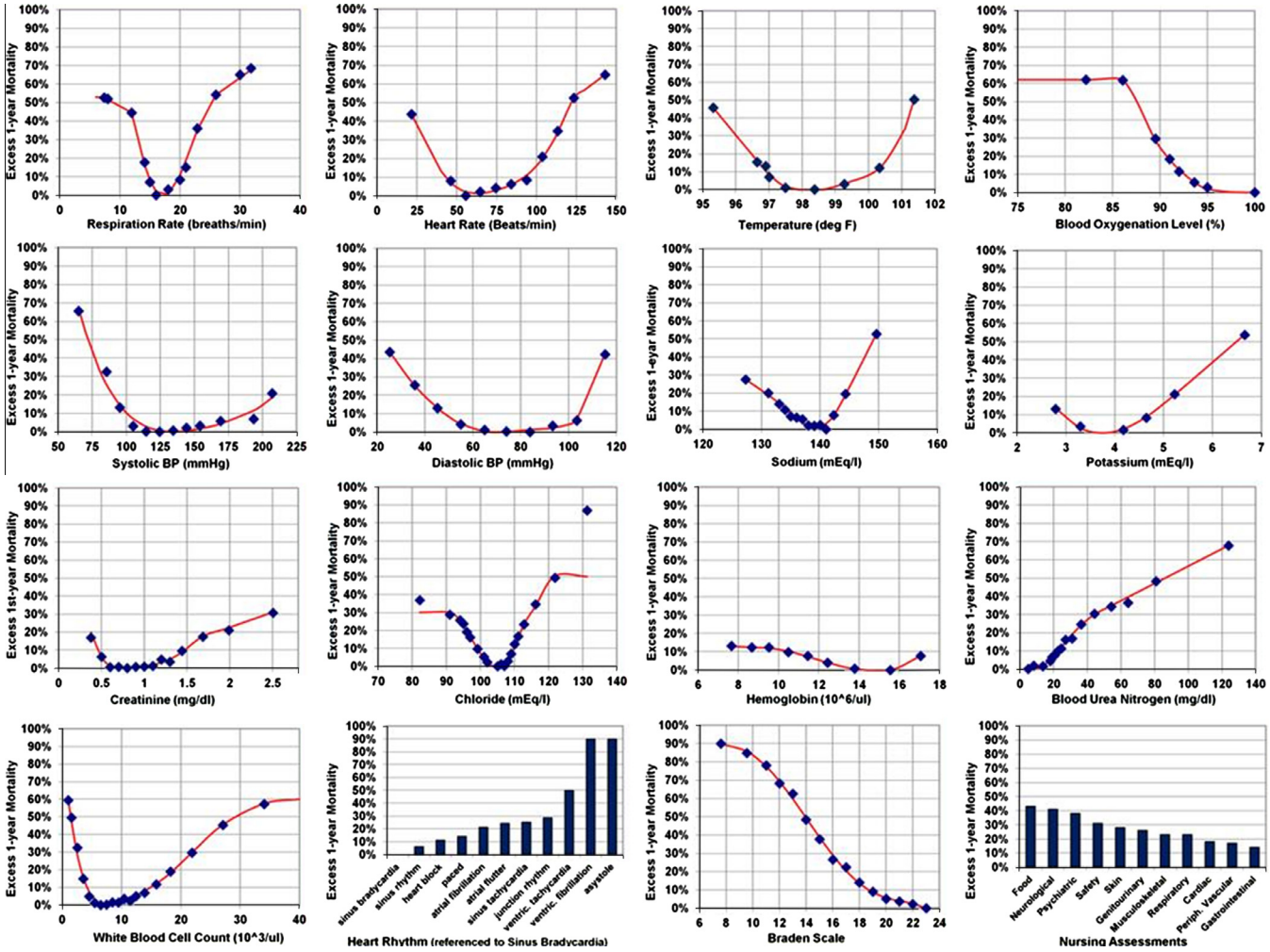

Fig. A1. Excess Risk Functions. In each plot, the final value of the variable prior to discharge is shown vs. 1-year all-cause mortality. Data are from 22,265 in-patients at SMH in 2004. Raw data are bucketed and a function is fitted to interpolate between bucket averages. Risk values are set to a constant above and below data extrema.

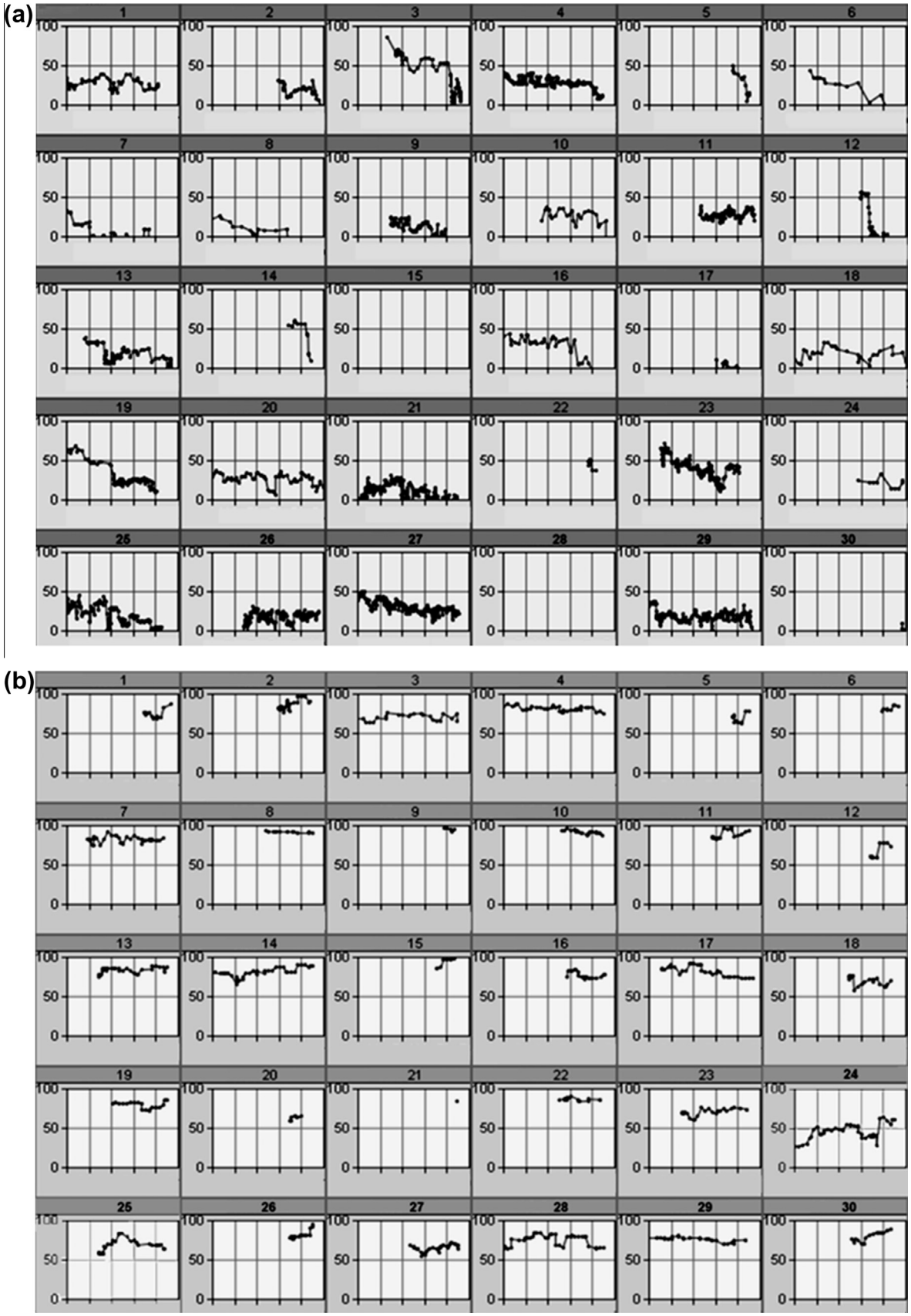

Fig. C2. Rothman Index graph arrays. (a) Graphs of 30 patients who died in the hospital. In a number of graphs the RI value drops below zero. (b) Graphs of 30 patients who were discharged to home. All graphs show the last 5 days of the patients' stay prior to discharge, and have an RI scale from 0 to 100. Graphs with a paucity of RI points indicate a short stay in the hospital or in (a) may be due to scores below zero.

To derive the smoothing function, we computed the ratio of RI_{noLab} over $RI_{withLab}$ for 34,612 instances from the SMH 2004 dataset in which new lab data had arrived and RI switched from RI_{noLab} to $RI_{withLab}$. We then grouped these ratios into 25 buckets, as a function of RI_{noLab}. For each bucket we computed the average ratio and the average RI_{noLab} and then fit a polynomial to the curve. This provided a variable scaling factor that is a function of overall risk. When RI_{noLab} outputs a high overall risk, the value of the smoothing factor is 0.9. In instances of very low risk the difference increases between the two sub-models increase and the smoothing factor falls below 0.6.

Appendix C. Applications of the Rothman Index: context and communication

Patient deterioration may occur hours or days in advance of critical events but can be difficult to discern [5,47,52]. For example, the National Patient Safety Agency reports that 11% of serious incidents are a function of deterioration not acted upon, and primary root causes include the failure to recognize the importance of a deterioration [53].

The challenges of early detection of patient deterioration are exacerbated by communication problems arising from fragmentation of care. The Institute for Healthcare Improvement observes that, with over 24 handoffs per patient during the course of an average 4.8-day stay, effective communication between caregivers is critical for maintaining continuity of care. [54] The Joint Commission reports that "80 percent of serious medical errors involve miscommunication between caregivers when patients are transferred or handed-off" and 37% of hand-offs are judged defective, in that the receiver does not have enough information to properly care for the patient. [55] Additionally, studies show that nurses deciding to call Medical Emergency Teams sometimes hesitate due to uncertainty in identifying changes in patient condition. [56] In any communication between caregivers, and especially at patient handoffs, information is lost, to the possible detriment of the patient [54,55,57]. Fortunately, the increasing use of information technology by clinicians, e.g. EMRs, can improve this situation. [58] The ability to see trends and changes in patient condition afforded by a continuously updated measure of patient condition may augment existing surveillance efforts [59] and is especially applicable to the "air traffic controller" model [51] where central monitoring can backstop clinicians who may not notice subtle patient deterioration over time.

The goal of this work was to develop and validate a continuous measure of general patient condition for medical-surgical as well as critical care patients using routinely recorded data from the Electronic Medical Record. Such a continuous measure of patient condition provides an opportunity to understand and communicate the current condition of a patient and may assist in the earlier detection of deterioration. By capturing, quantifying, and integrating data, the RI may provide a longitudinal perspective on both sudden and subtle changes to patient condition and assist in communicating these changes to clinicians.

Fig. A1 shows a graph of RI scores for a single patient over a 5-day period. Each point represents a new score computed when any of the 26 inputs was updated.

Presenting RI scores over time (e.g. graphically) provides context for understanding the implications of a patient's current condition and aids in early detection of deterioration. When viewing graphs of RI scores (Figs. C1 and C2):

- RI of 100 means that the patient is unimpaired.
- RI of 65 corresponds to the average acuity level of patients discharged to a SNF (see Table 3).
- RI of 40 corresponds to a MEWS score of 4, based on computations using the AMH dataset (using the MEWS model of Subbe et al.). A MEWS value of 4 is considered a critical score for activating an escalation pathway for intervention or transfer to a higher level of care [19,39].
- RI of 0 is generally the lowest score seen on a medical-surgical unit.
- Negative RI values are often seen in the ICU.

This RI graph (Fig. A1) illustrates potential opportunities for intervention throughout day 4, hours prior to the first Rapid Response Team call. Contextual understanding of patient condition as it is changing may improve the timeliness and confidence with which providers respond to deteriorating patients. This is particularly relevant when deciding to call Rapid Response and Medical Emergency Teams [48,56,60].

Fig. C1 shows RI graphs of patients in the final 5 days of their hospital stay. All the patients in Fig. C2a died in the hospital, while all the patients in Fig. C2b were discharged to home. Graphs were randomly sampled, subject to discharge criteria, from a common patient population (same 5-day date range from a single hospital).

A majority of graphs in Fig. C2a illustrate deterioration, as seen in the downward trend of RI, preceding death in accordance with clinical expectation. [61–63] Conversely, the trajectories of the patients discharged to home show predominantly flat or upward trends.

The impact of the RI on the effectiveness of communication and the quality of care will be the subject of future research [64,65].

References

[1] Adams ST, Leveson SH. Clinical prediction rules. BMJ 2012;344.
[2] Fraser DD, Singh RN, Frewen T. The PEWS score: potential calling criteria for critical care response teams in children's hospitals. J Crit Care 2006;21:278–9.
[3] Cuthbertson BH, Smith GB. A warning on early-warning scores! Br J Anaesth 2007;98:704–6.
[4] Liao L, Mark DB. Clinical prediction models: are we building better mousetraps? J Am Coll Cardiol 2003;42:851–3.
[5] Cuthbertson BH, Boroujerdi M, McKie L, et al. Can physiological variables and early warning scoring systems allow early recognition of the deteriorating surgical patient? Crit Care Med 2007;35:402–9.
[6] Cretikos M, Chen J, Hillman K, et al. The objective medical emergency team activation criteria: a case-control study. Resuscitation 2007;73:62–72.
[7] Devita MA, Bellomo R, Hillman K, et al. Findings of the first consensus conference on medical emergency teams. Crit Care Med 2006;34:2463–78.
[8] Gao H, McDonnell A, Harrison DA, et al. Systematic review and evaluation of physiological track and trigger warning systems for identifying at-risk patients on the ward. Intensive Care Med 2007;33:667–79.
[9] Hillman K, Chen J, Cretikos M, et al. Introduction of the medical emergency team (MET) system: a cluster-randomised controlled trial. Lancet 2005;365:2091–7.
[10] Subbe CP, Davies RG, Williams E, et al. Effect of introducing the Modified Early Warning score on clinical outcomes, cardio-pulmonary arrests and intensive care utilisation in acute medical admissions. Anaesthesia 2003;58:797–802.
[11] Winters BD, Pham J, Pronovost PJ. Rapid response teams–walk, don't run. JAMA 2006;296:1645–7.
[12] Egevad L, Granfors T, Karlberg L, et al. Prognostic value of the Gleason score in prostate cancer. BJU Int 2002;89:538–42.
[13] Berman M, Stamler A, Sahar G, et al. Validation of the 2000 Bernstein-Parsonnet score versus the EuroSCORE as a prognostic tool in cardiac surgery. Ann Thorac Surg 2006;81:537–40.
[14] Gogbashian A, Sedrakyan A, Treasure T. EuroSCORE: a systematic review of international performance. Eur J Cardiothorac Surg 2004;25:695–700.
[15] Pollack MM, Patel KM, Ruttimann UE. PRISM III: an updated pediatric risk of mortality score. Crit Care Med 1996;24:743–52.
[16] Rexius H, Brandrup-Wognsen G, Nilsson J, et al. A simple score to assess mortality risk in patients waiting for coronary artery bypass grafting. Ann Thorac Surg 2006;81:577–82.
[17] Knaus WA, Wagner DP, Draper EA, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. Chest 1991;100:1619–36.
[18] Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the working group on sepsis-related problems of the European society of intensive care medicine. Intensive Care Med 1996;22:707–10.

[19] Subbe CP, Kruger M, Rutherford P, et al. Validation of a modified Early Warning Score in medical admissions. QJM 2001;94:521–6.
[20] Kirkland LL, Malinchoc M, O'Byrne M, et al. A Clinical Deterioration Prediction Tool for Internal Medicine Patients. American Journal of Medical Quality Published Online First: 19 July 2012. doi: http://dx.doi.org/10.1177/1062860612450459.
[21] Escobar GJ, LaGuardia JC, Turk BJ, et al. Early detection of impending physiologic deterioration among patients who are not in intensive care: development of predictive models using data from an automated electronic medical record. J Hosp Med 2012;7:388–95.
[22] Prytherch DR, Smith GB, Schmidt PE, et al. ViEWS–Towards a national early warning score for detecting adult inpatient deterioration. Resuscitation 2010;81:932–7.
[23] Parshuram CS, Hutchison J, Middaugh K. Development and initial validation of the Bedside Paediatric Early Warning System score. Crit Care 2009;13:R135.
[24] McLellan MC, Connor JA. The Cardiac Children's Hospital Early Warning Score (C-CHEWS). J Pediatr Nurs Published Online First: 15 August 2012. doi: http://dx.doi.org/10.1016/j.pedn.2012.07.009.
[25] Buist MD, Moore GE, Bernard SA, et al. Effects of a medical emergency team on reduction of incidence of and mortality from unexpected cardiac arrests in hospital: preliminary study. BMJ 2002;324:387–90.
[26] Kyriacos U, Jelsma J, Jordan S. Monitoring vital signs using early warning scoring systems: a review of the literature. J Nurs Manag 2011;19:311–30.
[27] Duncan H, Hutchison J, Parshuram CS. The pediatric early warning system score: a severity of illness score to predict urgent medical need in hospitalized children. J Crit Care 2006;21:271–8.
[28] Mao Y, Chen Y, Hackmann G, et al. Early deterioration warning for hospitalized patients by mining clinical data. IJDKB 2011;2:1–20.
[29] Richardson DK, Gray JE, McCormick MC, et al. Score for neonatal acute physiology: a physiologic severity index for neonatal intensive care. Pediatrics 1993;91:617–23.
[30] Rothman MJ, Solinger AB, Rothman SI, et al. Clinical implications and validity of nursing assessments: a longitudinal measure of patient condition from analysis of the Electronic Medical Record. BMJ Open 2012;2. doi: http://dx.doi.org/10.1136/bmjopen-2012-000849.
[31] Walter LC, Brand RJ, Counsell SR, et al. Development and validation of a prognostic index for 1-year mortality in older adults after hospitalization. JAMA 2001;285:2987–94.
[32] Baid H. The process of conducting a physical assessment: a nursing perspective. Br J Nurs 2006;15:710–4.
[33] Kerr SD. A comparison of four nursing documentation systems. J Nurs Staff Dev 1992;8:27–31.
[34] Adams ST, Leveson SH. Clinical prediction rules. BMJ 2012;344:d8312.
[35] Boudreaux ED, Friedman J, Chansky ME, et al. Emergency department patient satisfaction: examining the role of acuity. Acad Emergy Med 2004;11:162–8.
[36] Grobman WA, Stamilio DM. Methods of clinical prediction. Am J Obstet Gynecol 2006;194:888–94.
[37] Lowry R. *Concepts & Applications of Inferential Statistics*, *Chapter 14*. <http://www.vassarstats.net/textbook> 2012.
[38] Kelly CA, Upex A, Bateman DN. Comparison of consciousness level assessment in the poisoned patient using the alert/verbal/painful/unresponsive scale and the Glasgow Coma Scale. Ann Emergy Med 2004;44:108–13.
[39] Higgins Y, Maries-Tillott C, Quinton S, et al. Promoting patient safety using an early warning scoring system. Nurs Stand 2008;22:35–40.
[40] Bland JM, Altman DG. Validating scales and indexes. BMJ 2002;324:606–7.
[41] Van Walraven C, Dhalla IA, Bell C, et al. Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community. CMAJ 2010;182(6).
[42] Lindenauer PK, Normand SL, Drye EE, et al. Development, validation, and results of a measure of 30-day readmission following hospitalization for pneumonia. J Hosp Med 2011;6(3):142–50.
[43] Hasan O, Meltzer DO, Shaykevich SA, et al. Hospital readmission in general medicine patients: a prediction model. J Gen Intern Med 2010;25(3):211–9.
[44] Hripcsak G, Albers DJ, Perotte A. Exploiting time in electronic health record correlations. J Am Med Inform Assoc 2011;18(Suppl 1):i109–15.
[45] Duckitt RW, Buxton-Thomas R, Walker J, et al. Worthing physiological scoring system: derivation and validation of a physiological early-warning system for medical admissions. An observational, population-based single-centre study. Br J Anaesth 2007;98:769–74.
[46] Inouye SK, Peduzzi PN, Robison JT, et al. Importance of functional measures in predicting mortality among older hospitalized patients. JAMA 1998;279:1187–93.
[47] Duncan KD, McMullan C, Mills BM. Early warning systems: the next level of rapid response. Nursing 2012;42:38–44.
[48] Odell M, Victor C, Oliver D. Nurses' role in detecting deterioration in ward patients: systematic literature review. J Adv Nurs 2009;65:1992–2006.
[49] Akre M, Finkelstein M, Erickson M, et al. Sensitivity of the pediatric early warning score to identify patient deterioration. Pediatrics 2010;125:e763–9.
[50] Smith GB, Prytherch DR, Schmidt PE, et al. Review and performance evaluation of aggregate weighted 'track and trigger' systems. Resuscitation 2008;77:170–9.
[51] Kho A, Rotz D, Alrahi K, et al. Utility of commonly captured data from an EHR to identify hospitalized patients at risk for clinical deterioration. AMIA Annu Symp Proc 2007;404–8.
[52] Liaw SY, Scherpbier A, Klainin-Yobas P, et al. A review of educational strategies to improve nurses' roles in recognizing and responding to deteriorating patients. Int Nurs Rev 2011;58:296–303.
[53] National Patient Safety Agency. Safer care for the acutely ill patient: learning from serious incidents. Available at: http://www.nrls.npsa.nhs.uk/resources/?EntryId45=59828; 2007.
[54] Institute for Healthcare Improvement: Improving transitions in hospital care. <http://www.ihi.org/offerings/VirtualPrograms/Expeditions/Transitions/Pages/default.aspx> (accessed 20 Feb2012).
[55] Joint Commission Center for Transforming Healthcare: Hand-Off Communications Project. <http://www.centerfortransforminghealthcare.org/projects/detail.aspx?Project=1> (accessed 20 Feb2012).
[56] Tait D. Nursing recognition and response to signs of clinical deterioration. Nurs Manag (Harrow) 2010;17:31–5.
[57] Blouin AS. Improving hand-off communications: new solutions for nurses. J Nurs Care Qual 2011;26:97–100.
[58] Bates DW, Gawande AA. Improving safety with information technology. N Engl J Med 2003;348:2526–34.
[59] Breslow MJ, Rosenfeld BA, Doerfler M, et al. Effect of a multiple-site intensive care unit telemedicine program on clinical and economic outcomes: an alternative paradigm for intensivist staffing. Crit Care Med 2004;32:31–8.
[60] Goldhill DR, McNarry AF. Physiological abnormalities in early warning scores are related to mortality in adult inpatients†. Br J Anaesth 2004;92:882–4.
[61] Cioffi J. Nurses' experiences of making decisions to call emergency assistance to their patients. J Adv Nurs 2000;32:108–14.
[62] Kause J, Smith G, Prytherch D, et al. A comparison of antecedents to cardiac arrests, deaths and emergency intensive care admissions in australia and New Zealand, and the United Kingdom—the ACADEMIA study. Resuscitation 2004;62:275–82.
[63] Hillman KM, Bristow PJ, Chey T, et al. Antecedents to hospital deaths. Int Med J 2001;31:343–8.
[64] Bradley EH, Yakusheva O, Horwitz LI, et al. Identifying patients at increased risk for unplanned readmission. Medical Care 2013, in press.
[65] Tepas III JJ, Rimar JM, Hsiao AL, Nussbaum MS. Automated analysis of electronic medical record data reflects the pathophysiology of surgical complications. Surgery 2013, in press.