Published in final edited form as:

J Crit Care. 2011 December ; 26(6): 634.e9–634.e15. doi:10.1016/j.jcrc.2011.05.003.

The use of an electronic medical record based automatic calculation tool to quantify risk of unplanned readmission to the intensive care unit: A validation study

Subhash Chandra, M.B.B.S.^{1,2}, Dipti Agarwal, M.B.B.S.¹, Andrew Hanson², Joseph C. Farmer, M.D.^{2,3}, Brian W Pickering, M.B., B.Ch.^{2,4}, Ognjen Gajic, M.D.^{2,3}, and Vitaly Herasevich, M.D., Ph.D.^{2,3}

¹ Department of Emergency Medicine, Mayo Clinic, Rochester, MN

² Multidisciplinary Epidemiology and Translational Research in Intensive Care (M.E.T.R.I.C.), Mayo Clinic, Rochester, MN

³ Division of Pulmonary and Critical Care Medicine, Department of Medicine, Mayo Clinic, Rochester MN

⁴ Department of Anesthesia, Mayo Clinic, Rochester MN

Abstract

Objective—To refine and validate an automatic risk of unplanned readmission (Stability and Workload Index for Transfer, SWIFT) calculator in a prospective cohort of consecutive medical ICU patients in a teaching hospital with comprehensive electronic medical records.

Design—Two phase (derivation and validation) prospective cohort study.

Settings—An academic medical ICU.

Subjects—Consecutive cohort of adult (age >18 years) patients with research authorization.

Intervention—The electronic medical record (EMR) based automatic SWIFT calculator.

Measurement—Agreement between the manual (“gold standard”) and automatic SWIFT calculation tool.

Main results—During the derivation phase, we enrolled 191 consecutive medical ICU patients. SWIFT scores for these patients calculated manually by the 2 reviewers had strong positive correlation, $r = 0.97$ and the mean (SD) difference was 0.43 (3.5). The first iteration of the automatic SWIFT calculator in the derivation cohort demonstrated excellent agreement with manual calculation, PaCO₂ ($k = 0.95$), PaO₂/FIO₂ ratio ($k = 0.69$), length of ICU stay ($k = 0.91$) and GCS ($k = 0.90$) and no agreement for source of ICU admission ($k = -0.15$). After adjustment in rules, the kappa value for hospital admission source improved to 1.0. Automatic calculation demonstrated strong correlation with manual, $r = 0.92$, and mean (SD) difference was -2.3 (5.9).

© 2011 Elsevier Inc. All rights reserved.

Correspondence Address: Vitaly Herasevich, Division of Pulmonary and Critical Care Medicine, Mayo Clinic, 200 1st St SW, Rochester, MN 55905, USA, herasevich.vitaly@mayo.edu, Tel.: +1-507-255-4055; Fax: +1-507-255-4267.

Institution: The work was performed at division of Pulmonary and Critical Care Medicine, College of Medicine, Mayo Clinic, Rochester, MN.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

During validation phase, 100 subjects were enrolled over 5 days. The automatic tool retained excellent correlation with gold standard calculation for SWIFT, $r = 0.92$, and the mean (SD) difference was -2.2 (5.5).

Conclusion—The EMR based automatic tool accurately calculates SWIFT score and can facilitate ICU discharge decisions without the need for manual data collection.

Keywords

Stability and Workload Index for Transfer; Electronic Medical Records; ICU readmissions

Introduction

Unplanned readmissions to intensive care units (ICU) are associated with significantly higher mortality and length of stay and the rate of ICU readmission is recognized as a quality indicator of performance. These readmissions are associated with significantly higher consumption of resources since their ICU length of stay is two times higher than new admissions, in-hospital mortality is 11 times higher, and overall mortality is up to six times higher. Recently our research team developed an index to facilitate the decision of patient discharge from an ICU, called the “Stability and Workload Index for Transfer (SWIFT).” This index was developed and validated from a prospective cohort of 1131 patients admitted to an academic medical ICU over a period of one year and relies on information readily available at time of discharge planning (e.g. ICU length of stay, patient admission source, Glasgow Coma Scale score, PaO₂/FIO₂, and the nursing demand for complex respiratory care). The index was further validated, externally. The overall discrimination power for ICU readmission was higher for SWIFT than APACHE III score when used at the time of ICU discharge.

An important barrier to the adoption of prediction indexes or scoring systems is the time taken to gather the required data to calculate specific clinical severity scores such as APACHE. Such data is routinely calculated by trained bedside nurses and requires over half an hour per patient. An additional factor limiting their use is the observed interobserver variability encountered; a significant problem even among experienced intensivists (when data is collected manually). Polderman et al. demonstrated that interobserver variability in data collection for APACHE II of 10–15% persists even with the use of strict guidelines and regular training programs. Computer algorithms apply the same rules or concepts each time and perform them with minimal input from bedside providers. Hence, developing algorithms which automatically generate scores such as APACHE or SWIFT may result in; 1) improvements in provider time usage; 2) decreased provider workload; 3) improved consistency of data collection with more accurate scores; and, 4) increased adoption (improved usability). The value of such algorithms in addressing these issues has been previously demonstrated. The implementation of electronic medical records (EMR) in critical care units facilitates the development and more routine use of these automatic tools.

This paper describes an EMR calculator intended to automatically calculate SWIFT scores on patients in an academic ICU setting equipped with a comprehensive EMR. The specific aim was to prospectively refine and validate the automatic SWIFT calculator in a cohort of consecutive medical ICU patients.

Methodology

Study Design and Settings

This prospective study was conducted in the 24-bed medical ICU of Saint Mary's Hospital, Mayo Medical Center, Rochester, MN. A comparative study design was used to compare the performance of an automatic SWIFT calculation tool with the gold standard manual SWIFT calculation as determined by two independent physician researchers. A two step process, derivation and validation, was used to define the performance of the automatic tool accurately. The study took place over 15 days in January 2010, (the first 10 days to recruit the derivation cohort and final 5 days to recruit the validation cohort.) The Institutional Review Board approved the study protocol and waived the need for informed consent.

Electronic Resources

The Critical Care Independent Multidisciplinary Program (IMP) of the study hospital has an established near-real-time relational database, the Multidisciplinary Epidemiology and Translational Research in Intensive Care (METRIC) Data Mart. Details of METRIC Data Mart and the EMR are previously published. METRIC Data Mart contains a near real time copy of patient's physiologic monitoring data, medication orders, laboratory and radiologic investigations, physician and nursing notes, respiratory therapy data, etc.. Access to the database is accomplished through open database connectivity (ODBC). Blumenthal et al. have reached on consensus through Delphi survey on defining a comprehensive EMR which requires 24 key functions to be present in all the clinical units of hospital. The current Mayo Clinic EMR fits the definition of a comprehensive EMR.

Subject Selection

A consecutive cohort of patients admitted to medical ICU was enrolled prospectively. Patients with age less than 18-years and those who denied research authorization were excluded from the study.

Data Collection and Score Calculation

Automatic SWIFT calculation- A software calculator was developed for automatic SWIFT score calculation. The components of SWIFT are described in Table 1. The SWIFT calculator is a SAS-based program which identifies current patients in the medical ICU each day at 6:45 AM and which then extracts those data points required for SWIFT calculation from patient demographics, physiological parameters and laboratory investigation domains of METRIC Data Mart. It uses the most recent available values of the score components at 6:45 AM. The calculated SWIFT score and corresponding component values from each run are stored in a database with a date/time stamp for further analysis.

“Gold standard” SWIFT calculation – two physician reviewers (SC, DA) collected individual SWIFT component values for all the patients present in the medical ICU at 6:45 AM, each day during the study period. The reviewers were blinded to the automatic tool results. The most recent value of SWIFT component available before 6:45AM was collected. Data was collected in the same manner during the validation phase. If a laboratory value (e.g., arterial blood gas) was not available for a given patient, then it was assumed to be normal/noncontributory. The value for each SWIFT component was allotted as originally described by Gajic et al. At the end of each phase, disagreements between the reviewers for each SWIFT component were resolved by consensus discussion. The final, manually collected, SWIFT score component values were defined as the “gold standard,” against which the performance of the automatic calculation tool was compared. The time required for manual data collection in each patient was recorded to the nearest one minute. All the disagreements between the manual and automatic calculation were reviewed by a physician

reviewer (SC). Based on the generated observations, the automatic calculation algorithm was iteratively modified in order to improve the accuracy of the SWIFT calculation. The newly modified algorithm was then re-run on the derivation cohort and the final algorithm was run on the validation cohort.

Outcomes

The primary outcome measured in this study was agreement between the manual (“gold standard”) and automatic SWIFT calculation tools. The agreement was determined for each SWIFT component as well as the calculated SWIFT score. The secondary outcome measure was the total time required to gather and calculate the SWIFT score for each individual patient while using the manual tool.

Statistical Considerations

Continuous variables are reported as mean, standard deviation (SD) and median and with interquartile range (IQR), based on distribution. Categorical data are reported as number and percent frequency of occurrence. The agreement for each SWIFT component was measured considering score allotted for each component as ordinal variable and reported as Cohen’s weighted kappa. The agreement in calculated SWIFT between automatic and gold standard calculation is reported using correlation coefficient and the difference is reported with SD. All the data analyses were performed using JMP statistical software.

Results

Derivation

Over the first 10 days (derivation phase), we screened 201 patients from medical ICU, 10 of them did not have research authorization and were excluded from the study. The clinical and demographic characteristics of the derivation cohort are shown in Table 2. The overall interobserver agreement between two reviewers for individual SWIFT components was excellent. The agreement was highest for PaCO₂, k (standard error) = 0.96 (0.013), followed by ICU length of stay, k (SE) = 0.95 (0.016) and least for source of ICU admission, k (SE) = 0.76 (0.037). When plotted using a Bland-Altman plot, SWIFT scores calculated manually by the 2 reviewers had very strong positive correlation, $r = 0.97$. The mean (SD) difference between two reviewers was 0.43 (3.5).

The first iteration of the automatic SWIFT calculator in the derivation cohort demonstrated excellent agreement with the gold standard on four out of five SWIFT components, k (SE) value of 0.69 (0.043) for PaO₂/FIO₂ ratio, 0.91 (0.028) for length of ICU stay, 0.90 (0.024) for GCS and 0.95 (0.029) for PaCO₂. The source of ICU admission captured using the automatic tool demonstrated no agreement with the gold standard, k (SE) = -0.15 (0.068). The calculated SWIFT demonstrated good correlation, $r = 0.92$. (Figure 1, Panel A) On the resulting Bland-Altman plot, the mean (SD) difference between the manual and automatic calculation was -3.5 (5.8), (Figure 2, panel A).

Rules adjustment

Bases on observations from derivation cohort study, the rules contained within the SWIFT score calculator algorithm were modified. These “new” rules extracted the source of ICU admission from the emergency department census table. Previously, this variable (admission source) was extracted from the billing information. With these modified rules, the automatic calculation was found to have 100% agreement with the gold standard for source of ICU admission. Calculated SWIFT score by these two methods again demonstrated strong correlation, $r = 0.92$. (Figure 1 panel B) The mean (SD) difference in calculated SWIFT score by two methods was -2.3 (5.9). In second run with revised rules, the difference in

calculated SWIFT via manual and automatic calculation decreases by 1.14, $p = 0.0296$. Manual calculation of SWIFT took an average of 2 minutes per subject, accounting for up to one hour per day of a trained person in a 24 bed medical ICU.

Validation

Over next 5 days, 100 subjects were enrolled in a validation cohort and all of these patients had approved research authorization (participation). The automatic tool retained excellent agreement with manual calculation for SWIFT components (Table 3). Calculated SWIFT score by two methods also maintained a strong correlation, $r = 0.93$ (Figure 2 panel C). The mean (SD) difference in calculated SWIFT score by two methods was -2.2 (5.5) (Figure 2, panel C).

Discussion

This study successfully derived and validated an automatic EMR-based tool for SWIFT score calculation in a medical ICU. The refined version of the automatic calculator required a revision of data set rules, based on observations made during the first run (derivation cohort). This revised tool had strong correlation with manual calculation, better agreement amongst SWIFT components, and significant decrease in difference in calculated SWIFT. This correlation persisted in the prospective validation cohort. In a 24 bed medical ICU, the automatic calculator has the potential to save one hour of manual work of an ICU care provider's time (usually a registered nurse).

The time and energy associated with the collection of data and calculation of commonly used ICU scores are often cited as major obstacles to the implementation of these sorts of tools. Like any other drug or device, the ready availability of clinical prediction models/indices often increase their use, and especially in a busy and time sensitive environment, such as an ICU. Hence, validated and reliable automatic calculators can play an important role in the implementation of new decision support tools.

The Institute of Medicine has suggested eight key components for safety, quality and care efficiency that should be included in an EMR. The first is to include computerized decision-support systems to prevent adverse drug interactions and improve compliance with "best practices." With expanding therapeutic options and increasing co-morbidities encountered in association with aging, both the demand for life support interventions and the complexity of critically ill patients are on the rise. This therapeutic intensity is reflected within EMR as an increase in the volume of documented information. A Canadian group of investigators estimated that the care of critically ill patients generates a median of 1348 individual data points/patient/day. This overload of information in EMR makes the manual data collection for calculation of a clinical prediction models more cumbersome. In a recent systematic review on barriers to the acceptance of EMR by physicians, Boonstra et al. classified the barriers into eight categories. One of the most reported categories was time; time required in implementation of the system, learning the system, to enter data, and time spent with each patient. Thus, the development and testing of automatic calculation tools decrease the time required for accurate data collection. Such systems may facilitate the adoption of EMRs which deliver meaningful use and decision support tools.

While the SWIFT calculation is simple enough to be performed by a care provider with little ICU background, it requires two minutes on average for each patient per day. For an individual patient this may be acceptable but for a 24 bed ICU, this imposes a significant burden, about 48 minutes to 1 hour each day in the study ICU. This is in addition to the potential for error associated with manual entry of data. The cost benefit analysis for implementing this automatic tool is beyond the scope of this paper but it seems clear that

with increased availability of sophisticated information technology in ICUs, simple, one time programming effort is less expensive than spending everyday almost one hour of an ICU care provider. The increasing availability of EMR in conjunction with the significant burden associated with manual collection will likely drive the development of automated alternatives such as the one presented here.

In addition to decreasing the workload on ICU care providers, computer-based automatic score calculators have been demonstrated to decrease the variability and errors associated with manual data collection. In the current study, the manual calculation in derivation cohort had 18% variability in calculated SWIFT score even after training and adherence to strict guidelines in a controlled setting. Multiple runs of the computer algorithm demonstrated 0% variability unless the logic rules used for SWIFT score calculation were changed. The automatic tool successfully eliminated the interobserver variability associated with manual calculation.

Strengths and Limitations

The prospective study design and strong infrastructure available at our institution enabled completeness of data. The data points for score calculation were obtained by two physician reviewers to have a robust and reliable reference standard to compare the performance of automatic tool. The major limitation of this study is the single ICU setting. Studies involving multiple ICUs are required to externally validate the performance of this automatic tool.

Conclusion

The presented study reports the development and validation of a fully automated calculation of the SWIFT score in a medical ICU. The developed tool demonstrated reliability and excellent correlation with the established gold standard manual calculation. Development of such automatic tools is helpful in decreasing the task load of bedside providers and reducing the error and variability in quality associated with manual data collection. Tested and validated automatic tools may facilitate the adoption of other predictive indexes such as Sequential Organ Failure Assessment (SOFA) and APACHE.

Acknowledgments

I acknowledge efforts of my co-authors for their invaluable contribution in completion of this study.

Financial support: This publication was made possible by grant 1 KL2 RR024151 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH), the NIH Roadmap for Medical Research, and Mayo Foundation. Its contents are solely the responsibility of the authors and do not necessarily represent the official view of the NCRR or NIH. Information on NCRR is available at <http://www.ncrr.nih.gov/>. Information on Reengineering the Clinical Research Enterprise can be obtained from <http://nihroadmap.nih.gov/clinicalresearch/overviewtranslational.asp>. This study was supported in part by National Heart, Lung and Blood Institute grant K23 HL78743-01A1 and NIH grant KL2 RR024151.

Figure 1.

Correlation between SWIFT score by manual calculation and automatic calculation. Panel-A, derivation cohort first run, $r = 0.92$; Panel-B, derivation cohort, second run, $r = 0.92$; Panel-C, validation cohort, $r = 0.93$ (SWIFT – stability and workload index for transfer)

Figure 2. Bland-Altman plot of calculated SWIFT (stability and workload index for transfer) score by manual calculation versus automatic calculation. Panel-A, derivation cohort first run; Panel-B, derivation cohort, second run; Panel-C, validation cohort.

Table 1
The Stability and Workload Index for Transfer (SWIFT) components and allotted score to individual components.

Variables	0	+1	+5	+6	+8	+10	+13	+14	+24
Original source of this ICU ^d admission	Emergency department				Transfer from a ward or outside hospital				
Total ICU length of stay	<2	2–10						>10	
Last measured PaO ₂ /FIO ₂ ^c ratio (during this ICU admission)	>400		<400 – >150			<150 – >100	<100		
GCS ^d at time of ICU discharge	>14			11 – 14				8 – 10	<8
Last arterial blood gas PaCO ₂ ^e (mm of Hg)	>45		<45						

^a ICU – Intensive Care Unit;

^b PaO₂- partial pressure of oxygen in arterial blood;

^c FiO₂ - fraction of inspired oxygen;

^d GCS – Glasgow comma scale;

^e PaCO₂ – partial pressure of carbon dioxide in arterial blood.

Table 2

Clinical and demographic characteristics of derivation and validation cohort

Variables	Derivation cohort (n = 191)	Validation Cohort (n = 100)
Age, median (IQR) ^a	65 (51 – 71)	64 (51 – 70)
Gender, female (%)	103 (54.8)	61 (61.0)
Source of ICU ^b admission		
ED ^c	79 (41.4)	35 (35.0)
Transfer	112 (58.6)	65 (65.0)
ICU length of stay, median (IQR)	1 (0 – 4) days	2 (0 – 6) days
GCS ^d , median (IQR)	14 (9 – 15)	14 (10 – 15)
SWIFT ^e , median (IQR)	18 (8 – 29)	14 (11–23)
Time required for SWIFT calculation/patient, mean (SD) ^f	1.7 (0.64) minutes	1.9 (0.62)

^aIQR – interquartile range;^bICU – Intensive Care Unit;^cED - emergency department;^dGCS – Glasgow comma scale;^eSWIFT – stability and workload index for transfer;^fSD – standard deviation.

Table 3

Interobserver agreement for SWIFT (stability and workload index for transfer) components between the gold standard and automatic calculation. Kappa and standard errors are presented for first run and run with modified rules on derivation cohort and run with modified rules on validation cohort.

SWIFT components	Agreement in first run (derivation cohort)	Agreement in second run (derivation cohort)	Validation cohort
Source of this ICU ^a admission	-0.15 (0.068)	1.00 (0)	1.00 (0)
ICU length of stay	0.91 (0.028)	0.91 (0.026)	0.93 (0.034)
Last measured PaO ₂ ^b /FIO ₂ ^c ratio	0.68 (0.043)	0.68 (0.043)	0.82 (0.049)
Glasgow Coma Scale	0.90 (0.024)	0.90 (0.024)	0.88 (0.041)
Last arterial blood gas PaCO ₂ ^d	0.95 (0.029)	0.95 (0.029)	0.91 (0.050)

^aICU – Intensive Care Unit;

^bPaO₂- partial pressure of oxygen in arterial blood;

^cFiO₂ - fraction of inspired oxygen;

^dPaCO₂ – partial pressure of carbon dioxide in arterial blood.

To calculate the agreement for each SWIFT component between two calculations, we considered score allotted for SWIFT component values as categorical values.