eHealth Beyond the Horizon – Get IT There
S.K. Andersen et al. (Eds.)
IOS Press, 2008

A Randomized Controlled Trial to Evaluate an Electronic Scoring Tool in the ICU

Thomas BÜRKLE[a1], Alexander BEISIG[a], Marion GANSLMAYER[b], Hans-Ulrich PROKOSCH[a]
[a]*Chair Medical Informatics, University Erlangen, Germany*
[b]*Department Internal Medicine 1, University Hospital Erlangen, Germany*

Abstract. Few RCTs on the effect of computer applications in intensive care have been published. This study presents an RCT measuring time savings and score values after introduction of a computer based scoring tool in an intensive care unit. A tablet PC with a standalone scoring application for TISS, SAPS2 and Apache 2 was supplied. We measured considerable time savings and higher score values when the computer application is used.

Keywords. RCT, Information system, Intensive Care

Introduction

When the influential 1999 IOM report "To Err is human: Building a Safer Health System" [1] was published, it prompted considerable discussion how better patient care could be achieved by using more computer applications. There have been critical reports however, that before stating that we could save so many lives we would need to thoroughly evaluate our computer applications and to prove that they fulfill the expectations [2]. However, it is still difficult to implement good evaluation studies for computer applications [3] and randomized controlled studies (RCTs) evaluating computerized information systems remain rare [4]. There have been many publications about electronic support and patient data management systems (PDMS) in intensive care units ICU [e.g. 5,6] but only few RCT's on this topic [7,8,9,10]. A Medline search refined for RCTs with the MeSH terms "Information Systems" and "Intensive Care" revealed only six publications.

This paper presents an RCT at the medical intensive care unit 1 (MICU-1) of Erlangen University Hospital, which was performed to evaluate a standalone scoring application on a tablet PC. Our goals was to assess user satisfaction (by means of a questionnaire), to find out how workflows changed (by workflow analysis), to measure time needed to score a patient (by RCT) and to compare score values generated with help of computer and without (by RCT).

[1] Corresponding Author: PD. Dr. Thomas Bürkle, Lehrstuhl für Med. Informatik, Krankenhausstrasse 12, D 91054 Erlangen, Germany, Email: Thomas.Buerkle@imi.med.uni-erlangen.de

1. Environment

Our university hospital has several ICUs. MICU-1 belongs to the department Internal Medicine 1 which specializes for gastroenterology, pneumology and endocrinology. MICU-1 has 12 beds, 44 nurses and 11 physicians in rotational shifts. In 2005 they cared for 537 patients and delivered more than 48.000 respirator hours and 499 days of CVVH. Therapy comprises diseases like ARDS, pneumonia, sepsis, liver failure, gastrointestinal bleeding, severe clotting disorder, poisoning, metabolic failure as well as post-transplantation complications, usually combined with multi organ failure. In contrary to some other Erlangen University Hospital ICUs there is currently no PDMS available on MICU-1. At the time of this study the computing infrastructure of MICU-1 comprised four workstations in the physician office and two more in the nursing office to access e.g. the Erlangen Hospital Information System HIS (Siemens Soarian) and to perform lab order entry (Swisslab) and radiology order entry (iSoft RadCentre). The workstations in the nursing office support electronic drug ordering from the pharmacy. There is a central monitoring system for all beds at the nursing office (Siemens Infinity). Furthermore a PACS workstation and printer are available for retrieval of radiology images (Siemens). The patient chart and patient record however are maintained completely on paper.

German legal and reimbursement requirements [11] necessitate since 2006, that for each intensive care patient the scores CORE10TISS, a TISS28 derivate (therapeutic intervention scoring system [12,13]) and SAPS2 (simplified acute physiology score [14]) are collected daily. Hospitals must supply accumulated score sums for each ICU patient. The totaled score is mapped to the German OPS-Code (an ICPM derivate) 8-980 which stands for complex intensive care treatment. In combination with the totaled respirator hours this data is used for reimbursement. Traditionally and for scientific evaluations MICU-1 collects daily two other scores for each patient, namely the APACHE2 [15] and the TISS76 [16] score. Before this study, all scoring was done manually on paper and the results entered in the paper based patient chart and a small MS-Access database application for statistical evaluation. This was a time consuming task.

2. Methods

On request, the medical informatics department analyzed the current situation and documentation workflow at MICU-1 and offered to implement a computerized scoring tool as an interim solution as long as no PDMS is available. The application was designed for tablet PC to support bedside score documentation and programmed in JAVA using Hibernate 3.2 and HSQL 1.8.0 as database and the Apache FOP library for graphical output. It's specifications comprised download compatibility into the existing Access database, easy configuration of new scores with XML template, and user profiling to permit each user to define his own sequence of data entry. Application development was based on rapid prototyping in close collaboration with clinicians. The scoring tool went into use in December 2006. Figure 1 shows the user interface which is optimized for pen entry.

To evaluate the new scoring tool a complex evaluation protocol was defined combining workflow analysis, time series questionnaire technology and a randomized trial measuring time consumption and score values. Here we will concentrate on the

RCT results to answer the hypotheses “the system has changed the time needed to document scores” and “the system has changed the quality of score documentation”. To support or reject these hypotheses n=54 patient scoring events performed in March and April 2007 have been groupwise randomized into 27 interventions where the computer application was used for scoring and 27 controls where scoring was performed manually as before. A scoring event is defined as collecting four scores CORE10TISS, SAPS2, APACHE2 and TISS76 once at midnight for one patient. Time measurement was done manually using a stop watch. Measurement started when the paper based patient chart (plus tablet PC in intervention) were available and finished when the respective score values were written in the patient chart. In the manual group each of the four scores was timed separately, whereas in the intervention group we measured time for SAPS2 and APACHE2 combined, because the documentation workflow had been streamlined in the computer application. There, physiology parameters for both SAPS2 and APACHE2 could be documented just once and the program did allocate the respective score values for both (see Figure 1). For total time measurement, times to fetch or distribute the paper patient records were included.

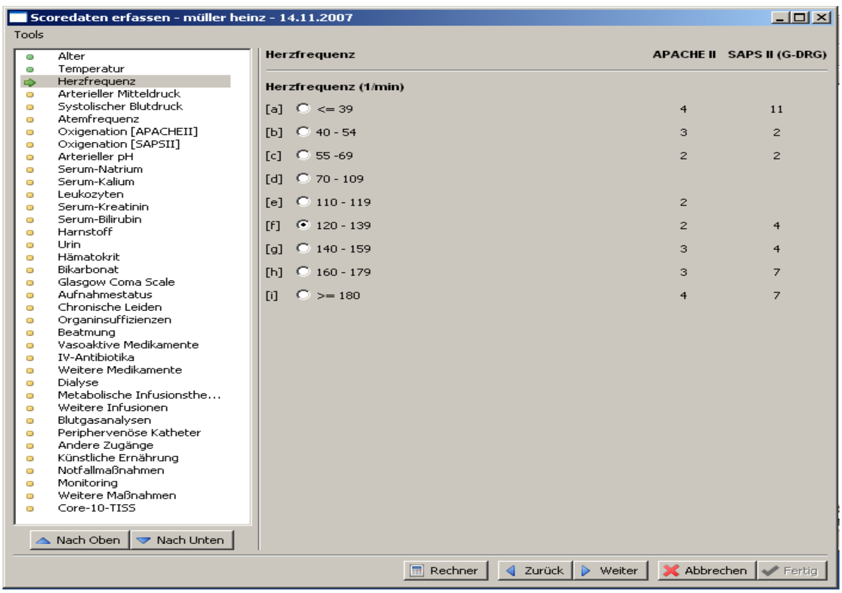

Figure 1: data entry screen of score tool on tablet PC

Furthermore we collected all score values from intervention and control group. In order to assess the quality of the documented score values, all scores of both groups have been collected a second time from the researcher (author AB) using the computer tool. Then measurements have been collected a third time from a senior physician of the ward, thus establishing the presumably correct gold standard.

For statistic testing we used Mann Whitney U test assuming 5% error level.

3. Results

Median values of scoring time measurements in seconds are shown in Figure 2. Documentation time for the combination of SAPS2 and APACHE2 (both times have been summed in the manual group) is nearly halved from 150 to 80 seconds with computer assistance. For TISS76 the time saving is 60 seconds on computer compared to 95 seconds on paper, for CORE10TISS it is 10 against 15 seconds and the total time is 165 seconds against 265 seconds. All time differences are significant at the 5% level.

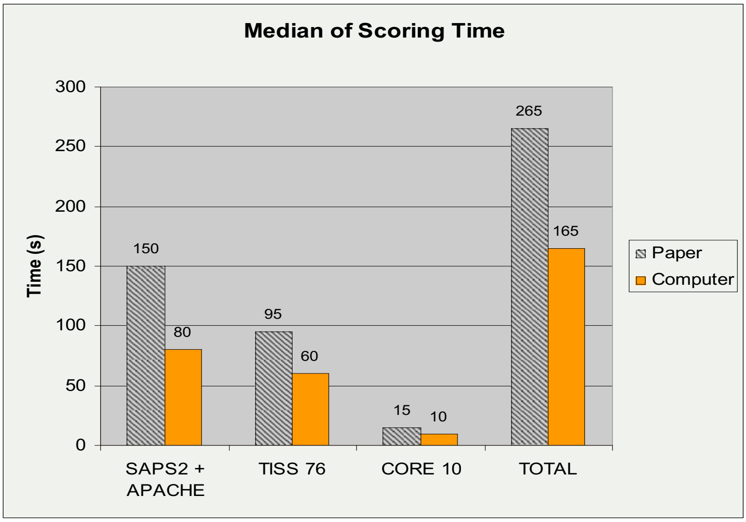

Figure 2. Median time values for scoring using paper or computer

Figure 3 shows score results for APACHE paper versus gold standard on the left (27 randomized score values measured twice) and scoring tool versus gold standard (another 27 randomized score values measured twice).

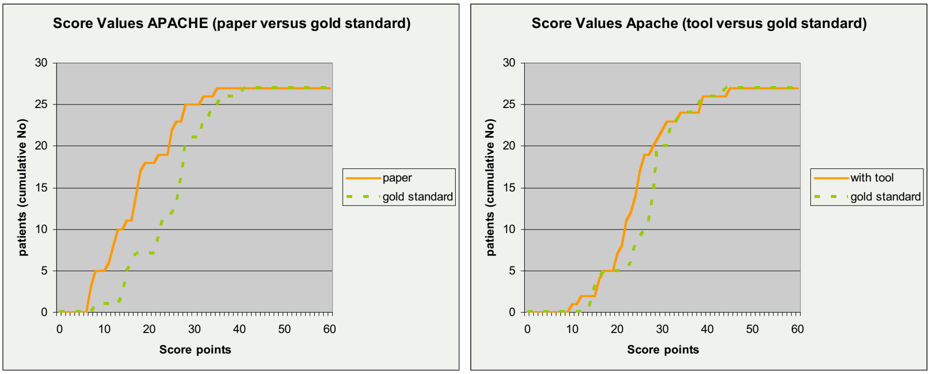

Figure 3. Score values compared to gold standard, control (paper based scoring) left, intervention (computer tool) right

APACHE2 scores on paper (Figure 3 left) were 10 points lower than gold standard (significant at 5%), compared to scores in the intervention group with computer (Figure 3 right) which were only 4 points lower than gold standard (non significant, p=6.4%). Accordingly we found a difference of 6 points to the gold standard for SAPS2 when using paper (significant) compared to only 4 points difference when using computer (non significant). For TISS76, values using paper were 4 points lower than gold standard (significant) while computer score values were 2 points higher than gold standard (non significant). For CORE10TISS there were no significant differences between measurement and gold standard neither in intervention nor in control group. In summary we note a significant decrease of scoring time when using computer and higher score values near to our gold standard for scoring APACHE2, SAPS2 and TISS76 on computer.

4. Discussion

Within this evaluation we could demonstrate a significant decrease in time used for scoring intensive care patients by introducing a plain computer based scoring tool. These results correspond nicely with the results of our user satisfaction evaluation which is beyond the scope of this paper. We identified one other RCT [7] which did also demonstrate time savings for computerized documentation, but did so for a full grown commercial PDMS and measured nursing activities. We measured physicians work time and our intervention was only a minor stand alone application taking less than 3 man months for development.

Furthermore we could demonstrate a significant influence of computerized scoring on scoring quality, since manual score values for three out of four scores were too low. It remains difficult to establish a valid gold standard for scoring ICU patients. One physician may e.g. consider a documented high pulse or pressure value from a monitor as an artefact due to patient movements whereas another colleague may trust this value because he had attended the measurement himself and therefore is confident that the peak reflects reality. Therefore our score values are not fully identical between gold standard (established by repeated measurements) and both computer scoring or paper scoring. It remains a fact however, that those differences are small and insignificant when computerized scoring was used, whereas values were significantly lower when paper was used, except for the elementary CORE10TISS score. There is one other study of Bosman and colleagues [17] who performed simultaneous manual and computerized charting using a PDMS and note, similar to our study, higher severity scores for APACHE2, SAPS2 and MPM (Mortality Probability Models) resulting in an increased predicted mortality. They do however not use a gold standard such as repeated scoring to confirm which values are better. In that study, physiology parameters were drawn directly from the PDMS for computerized scoring and from the paper record for manual scoring, whereas our source information was always a paper based patient record. Therefore we conclude that even without direct and automated transmission of vital signs into an electronic patient chart a computerized documentation tool may alter scoring results.

However, critique to this evaluation study should be allowed. This is a small study with n=54 scoring events. The designer of the application took part in the evaluation of his own program, which could lead to a Hawthorne effect. In a disruptive environment such as an intensive care unit, there are many interfering factors which may influence an interventional study. We tried to overcome both facts by hopefully distributing interfering factors equally among intervention and control group using a rigid study protocol and by measuring differences against a gold standard.

Our results suggest that ICU scoring of non trivial scores should always be done with computer support, not only to save time, but to achieve a higher quality of score values. Considering the scoring application itself we are fully aware that we have dealt only with one minor documentation aspect on MICU-1. We plan to support this ward in future with a commercial PDMS which will draw vital signs for scoring SAPS2 directly from imported patient monitor data. But our results indicate, that we should increase computerized documentation in ICUs for better data quality, despite of the risks which arise from increased dependency of computer applications.

References

[1] L. Kohn, J. Corrigan, M. Donaldson (eds), *To Err is Human: Building a Safer Health System.* Washington DC: National Academic Press, 1999.

[2] Institute for Safe Medication Practices, Over-reliance on computer systems may place patients at great risk. *ISMP Medication Safety Alert* Huntingdon Valley, Pa.: ISMP,1999.

[3] T. Bürkle, E. Ammenwerth, H.U. Prokosch, J. Dudeck, Evaluation of clinical Information Systems - What can be evaluated and what cannot ? *J Eval Clin Pract.* **7(4)** (2001), 373-85.

[4] E. Balas, S. Austin, J. Mitchel, B. Ewigmann, K. Bopp, G. Brown, The clinical value of computerized information services. A review of 98 randomized clinical trials. *Archives of Family Medicine* **5** (1996), 271-278.

[5] T. Bürkle, A. Michel, W. Horch, L. Schleifenbaum, J. Dudeck, Computer based Nursing Documentation - Means to achieve the goal. *Int J Med Inf.* **52(1-3)** (1998), 71-80.

[6] R. Röhrig, S. Böttger, A. Junger, M. Benson, A. Michel, G. Hempelmann, Vollautomatische Score-Evaluierung mit einem Patienten-Daten-Management-System (PDMS). *Anästhesiologie & Intensivmedizin* **42** (2001), 409-477.

[7] R. Bosman, E. Rood, H. Oudemans-van Straten, J. Van der Spoel, J. Wester, D. Zandstra, Intensive care information system reduces documentation time of the nurses after cardiothoracic surgery. *Intensive Care Med* **29** (2003), 83-90.

[8] E. Rood, R. Bosman, J. van der Spoel, P. Taylor, D. Zandstra, Use of a computerized guideline for glucose regulation in the intensive care unit improved both guideline adherence and glucose regulation. *JAMIA* **12,2** (2005), 172-180.

[9] S. Weingarten, M. Redinger, L. Conner, T. Lee, I. Hoffman, B. Johnson, G. Ellrodt, Practice guidelines and reminders to reduce duration of hospital stay for patients with chest pain: an interventional trial. *Annals of Internal Medicine* **120(4)** (1994), 257-263.

[10] K. Colpaert, B. Claus, A. Somers, K. Vandewoude, H. Robays, J. Decruyenaere, Impact of computerized physician order entry on medication prescription errors in the intensive care unit: a controlled cross sectional trial. *Critical Care* **10(1)** (2006), R21

[11] B. Rapp, DRG-Update Intensivmedizin 2006. *Das Krankenhaus* **1** (2006), 32-39.

[12] D. Cullen, M. Civetta, B. Briggs, L. Ferrara, Therapeutic intervention scoring system: a method for quantitative comparison of patient care. *Crit Care Med* **2** (1974), 57-60.

[13] D. Miranda, A. de Rijk, W. Schaufeli, Simplified Therapeutic Intervention Scoring System: the TISS-28 items--results from a multicenter study. *Crit Care Med.* **24(1)** (1996), 64-73.

[14] J. Le Gall, S. Lemeshow, F. Saulnier, A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA* **270(24)** (1993) 2957-2963.

[15] W. Knaus, E. Draper, D. Wagner, J. Zimmerman, APACHE II: A severity of disease classification system. *Crit Care Med* **13** (1985), 818-829.

[16] R. Keene, D. Cullen, Therapeutic Intervention Scoring System Update 1983. *Crit Care Med* **11** (1983), 1-3.

[17] R. Bosman, H. Oudemans-van Straten, D.F. Zandstra. The use of Intensive care information systems alters outcome prediction. *Intensive Care Med* **24** (1998), 953-58.