Artificial Intelligence in Medicine 92 (2018) 43–50

Contents lists available at ScienceDirect

Artificial Intelligence in Medicine

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

Using Arden Syntax Medical Logic Modules to reduce overutilization of laboratory tests for detection of bacterial infections—Success or failure?

Ixchel Castellanos[a,*,1], Stefan Kraus[b,c,1], Dennis Toddenroth[c], Hans-Ulrich Prokosch[b,c], Thomas Bürkle[d]

[a] Department of Anaesthesiology, University Hospital Erlangen, Krankenhausstr. 12, 91054 Erlangen, Germany
[b] Center for Communication and Information Technology, University Hospital Erlangen, Glückstr. 11, 91054 Erlangen, Germany
[c] Department of Medical Informatics, Biometrics and Epidemiology, Chair of Medical Informatics, Friedrich-Alexander-University Erlangen-Nuremberg, Wetterkreuz 13, 91058 Erlangen, Germany
[d] Bern University of Applied Sciences, Institute for Medical Informatics, Höheweg 80, CH-2502 Biel, Switzerland

ARTICLE INFO

Article history:
Received 14 January 2015
Received in revised form 12 September 2015
Accepted 12 September 2015

Keywords:
Clinical decision support system
Procalcitonin
Overutilization
Underutilization
Arden Syntax

ABSTRACT

Objective: Bacterial infections frequently cause prolonged intensive care unit (ICU) stays. Repeated measurements of the procalcitonin (PCT) biomarker are typically used for early detection and follow up of bacterial infections and sepsis, but those PCT measurements are costly. To avoid overutilization, we developed and evaluated a clinical decision support system (CDSS) in Arden Syntax which computes necessary and preventable PCT orders.
Methods: The CDSS implements a rule set based on the latest PCT value, the time period since this measurement, and the PCT trend scenario. We assessed the CDSS effects on the daily rate of ordered PCT tests within a prospective study having two ON and two OFF phases in a surgical ICU. In addition, we performed interviews with the participating physicians to investigate their experience with the CDSS advice.
Results: Prior to the deployment of the CDSS, 22% of the performed PCT tests were potentially preventable according to the rule set. During the first ON phase the daily rate of ordered PCT tests per patient decreased significantly from 0.807 to 0.662. In subsequent OFF, ON and OFF phases, however, PCT utilization reached again daily rates of 0.733, 0.803, and 0.792, respectively. The interviews demonstrated that the physicians were aware of the problem of PCT overutilization, which they primarily attributed to acute time constraints. The responders assumed that the majority of preventable measurements are indiscriminately ordered for patients during longer ICU stays.
Conclusion: We observed an 18% reduction of PCT tests within the first four weeks of CDSS support in the investigated ICU. This reduction may have been influenced by raised awareness of the overutilization problem; the extent of this influence cannot be determined in our study design. No reduction of PCT tests could be observed during the second ON phase. The physician interviews indicated that time critical ICU situations can prevent extensive reflection about the necessity of individual tests. In order to achieve an enduring effect on PCT utilization, we will have to proceed to electronic order entry.

1. Background

1.1. Scientific background

In intensive care environments, sepsis and the systemic inflammatory response to infectious agents is a frequent cause of severe clinical outcomes [1] with an incidence of 75,000 cases yearly and a hospital mortality rate of 55% in Germany in 2007 [2]. For Germany, it has been estimated that 30% of the intensive care unit (ICU) costs are caused by severe sepsis [3], totaling 1.77 billion

* Corresponding author. Tel.: +49 9131 85 33680; fax: +49 9131 85 32907.
E-mail address: ixchel.castellanos@kfa.imed.uni-erlangen.de (I. Castellanos).
[1] Contributed equally.

https://doi.org/10.1016/j.artmed.2015.09.005
0933-3657/

Euros per year. Sepsis can be difficult to detect because symptoms are polymorphic and frequently superimposed on the patient's underlying disease [4], but rapid diagnosis and antibiotic therapy determines treatment success [5]. Previously, laboratory tests concentrated on leucocytosis or C-reactive protein levels (CRP) to recognize the accompanying inflammatory response, however today the biomarker procalcitonin (PCT) is increasingly used for sepsis detection in ICU environments due to its faster response and the ability to discriminate between an infectious bacterial or non-bacterial etiology [6–9]. ICU monitoring of sepsis with PCT tests requires repeated tests to investigate the progression of the inflammatory reaction and the effectiveness of therapy.

1.2. Rationale for the study

Several examples of excessive utilization of laboratory tests, termed *overutilization*, have been reported in the literature [10,11], and a variety of countermeasures has been devised, for an overview see [12]. Recommended countermeasures comprise "education" and "raising awareness" as well as the use of clinical decision support systems (CDSS) [13]. This investigation was prompted by the observation that routinely requested PCT measurements at a local surgical ICU had a tendency to substantially exceed clinical needs. An initial estimation showed that about 15% of the PCT tests with redundant costs of 14 euro per test might be preventable.

1.3. Objectives of the study

The objective of this study was the development and evaluation of a knowledge module to assist physicians in deciding whether PCT tests for an individual patient are required. The evaluation assessed quantitative effects of the CDSS on the daily rate of ordered PCT measurements and was accompanied by a qualitative component to examine the acceptance of the CDSS among physicians.

2. Study context

2.1. Organizational setting

Erlangen University Hospital is a tertiary care university hospital with 1361 beds in 2012. In 2006, we introduced a commercial Patient Data Management System (PDMS), which was initially deployed in a large surgical ICU (SICU) [14]. Today nine ICUs with a total of approximately 130 beds have been PDMS equipped. This study took place on the pilot SICU with 25 beds and a census of approximately 2000 patients per year. The SICU is divided into a general surgery section with 16 beds and a cardiac surgery section with 9 beds.

2.2. System details and system in use

Within the SICU the PDMS covers complete bedside documentation and some computerized physician order entry (CPOE) functionalities. All vital signs, medication orders, medical and nursing activities are captured at bedside workstations and collected into a digital patient chart. This includes daily patient Simplified Acute Physiology Score (SAPS II) and Acute Physiology And Chronic Health Evaluation (APACHE) scoring. The PDMS is interfaced with medical devices such as patient monitors and ventilators and with the hospital information system, including import of laboratory results, microbiology, radiology reports, diagnoses and procedures as well as export of Diagnosis Related Groups (DRG) relevant data [15]. Currently, CPOE includes external orders such as laboratory requests, but is performed using separate software based on a special lab list which is filled out by the attending physician.

In a cooperation project with the system vendor, the PDMS (Integrated Care Manager, ICM®, Dräger Medical, Lübeck, Germany) has been enhanced with modular decision support functions based on the Arden Syntax for Medical Logic Systems. This language has been designed to represent medical knowledge in a standardized and shareable form [16], and has found its traditional role for data-driven clinical event monitoring [17–19]. Arden Syntax "may be considered as a hybrid between classical production rules and procedural representation of clinical algorithms" [20]. In Arden, medical knowledge is contained in Medical Logic Modules (MLMs). The process of CDSS integration with the PDMS at our institution has been described in a previous paper [21]. The resulting system offers some features beyond the typical capabilities of embedded Arden-based systems [22]. In particular, our CDSS is capable of immediate user-driven decision support at the point of care based on an MLM viewer which has been integrated into the PDMS GUI. Another feature is the ability to provide list-based multi-patient decision support, whereas most Arden systems described in literature act only in a single patient context [23]. Our integration was based on a commercial Arden engine (ArdenSuite®, Medexter Healthcare, Vienna, Austria) supporting all versions of Arden Syntax. Access to the PDMS patient database has been implemented with a periodic replication job which feeds the CDSS facts database, because the PDMS does not permit direct data access by third party systems.

3. Methods

3.1. Study design

We implemented a longitudinal prospective study design covering all surgical ICU patients and physicians in charge within the study period. The study was implemented as an ON–OFF–ON–OFF design (turning the CDSS on and off twice) without randomization. The intervention comprised the activation of a PCT MLM which generates individualized recommendations whether a PCT test for a certain patient should be performed or not. An initial instruction of all participating physicians took place. The MLM thereby produces a list of suggested PCT tests for both the general and the cardiac surgery section of the SICU. The user-driven PCT MLM had to be invoked intentionally by the physician using a button for each of the sections within the PDMS. It returns a list of all ICU patients with individual PCT recommendations and a short explanatory statement for each patient (Fig. 1).

During both ON phases the MLM was activated automatically using the temporal operators of the Arden Syntax. During the OFF periods the PCT MLM informed the user that currently no decision support was available. The physicians of the early shifts, when PCT ordering was performed, were repeatedly encouraged by the senior consultant to use the PCT MLM during ON phases.

We calculated a baseline estimate of preventable PCT tests by retrospectively applying the MLM rule set to the tests that were ordered within four months before the study started (PRE phase). A terminal POST phase was added to examine the further progression of daily PCT rates without CDSS. The study was followed by a series of semi-structured interviews with the ICU physicians to determine reasons for use or disuse of the PCT MLM.

3.2. Background

In the examined SICU the ordering process for PCT tests is still partially paper-based. The physician triggers laboratory tests using a paper document called "special lab orders list" which also comprises various extra order options for all current ICU patients. Each physician fills out the special lab orders list for the patients in

Bed / Patient	Recommendation
Bed 09 (4/1) #anonymized#	**Recommendation:** Do not measure PCT now **Rationale:** The latest PCT measurement is not older than 36 hours (6 hours 22 minutes) The latest PCT measurement (0.38 ng/ml) does not exceed the threshold of 1 ng/ml The current PCT trend (-0.06 ng/ml/24h) does not exceed the threshold of 0.20 ng/ml/24h
Bett 10 (4/2) #anonymized#	**Recommendation:** Measure PCT now **Rationale:** The latest PCT measurement (3.19 ng/ml) exceeds the threshold of 1 ng/ml
Bett 11 (4/3) #anonymized#	Criterion cannot be assessed, because no PCT values are available so far

Fig. 1. Detail of the recommendations list for three patients. The first message formatted in green indicates that the patient did not require a PCT test. The second message formatted in red indicates that the patient should have a PCT test. The third message formatted in black indicates that there are no previous PCT values, therefore no recommendation can be given for that patient (PCT = procalcitonin, ng = nanograms, ml = milliliters, h = hours). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

her/his responsibility at 8 a.m. on weekdays and 2 a.m. on weekends. The special lab orders list contains a "Sepsis" box which may be ticked. A ward clerk takes the list and transfers it into computerized lab orders outside the PDMS, e.g. a PCT order if "Sepsis" has been ticked. The laboratory results for all lab parameters are returned directly into the PDMS via an HL7 interface.

For the PCT MLM we created a rule set for required and preventable PCT tests based on conservative assumptions with priority on patient safety. Fig. 2 demonstrates a typical pattern of PCT progression which is available within the CDSS component of the PDMS. Initially, the patient shows indications of a fresh inflammatory response, resulting in a high PCT measurement frequency. Circles comprise all PCT exams which have been ordered for the patient. Those four PCT exams which would have been considered avoidable according to the MLM rule set are marked by black circles and vertical lines.

The rule set constitutes the centerpiece of the PCT MLM and has been devised by the senior consultant. It comprises just three rules. If any of these conditions is true, a PCT test is recommended:

(1) Latest PCT value above a threshold (1 ng/dl).
(2) Latest PCT value older than a time limit (36 h).
(3) Current increase between latest two PCT values above a trend limit (0.2 ng/dl/24 h).

The third condition requires at least two consecutive PCT values. If there are not enough values for a particular patient, the PCT MLM acts cautious stating that for this patient a recommendation is unavailable (Fig. 1).

3.3. Participants

Participants of the study were thirteen physicians working at any of the two sections of the SICU during any of the four ON and OFF phases. There were no additional selection criteria. No randomization was applied in order to prevent compromising patient safety.

The initial instruction of all participating physicians consisted of an email informing about the upcoming study approximately two weeks before activating the MLM and a reminder email one week before activating the MLM. At day one of phase ON1 the use of the MLM was demonstrated by the senior physician to all physicians in that shift in a short instruction (approx. 5–10 min). The instruction explained:

- the background of the study ("to avoid over- and underutilization of PCT testing"),
- where to find and how to start the CDSS (if not already known from earlier use of other MLMs),
- how to start the PCT MLM,
- the three possible outputs ("measure", "do not measure", "no criteria"),
- that the MLM should be used every morning before filling out the special lab list,
- that the physicians are always free to follow the recommendation of the CDSS or to ignore it and follow their clinical experience and intuition about the necessity of PCT testing in a certain case.

During the first five days of both ON phases, the physicians of the morning shift were reminded to use the CDSS.

For the interviews we contacted all participants of the study. One had already left our institution and was no longer available. Six out of the remaining twelve participants consented to be interviewed.

3.4. Study flow

The study consisted of six phases as shown in Fig. 3. During PRE (four months), we analyzed PCT measurements by calculating the percentage of PCT tests which where preventable according to the rule set in order to determine the required sample size for the intervention. During this period, physicians were not yet informed about the forthcoming study in order to avoid any bias that could

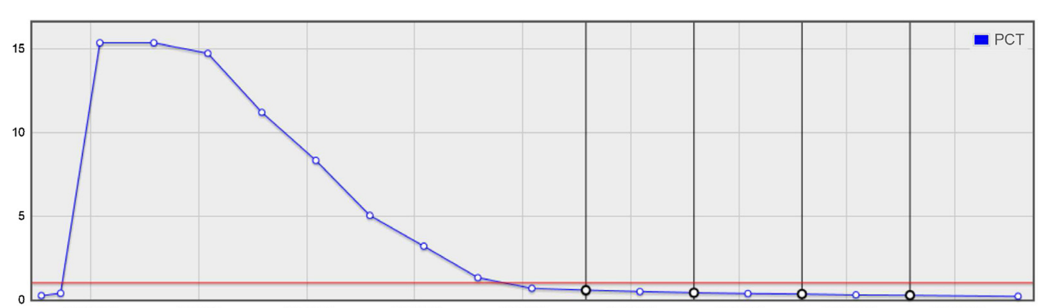


Fig. 2. Typical pattern of PCT progression. Each ordered PCT test is marked by a circle. Four black circles with vertical lines mark PCT tests which could have been omitted according to the rule set (PCT = procalcitonin).

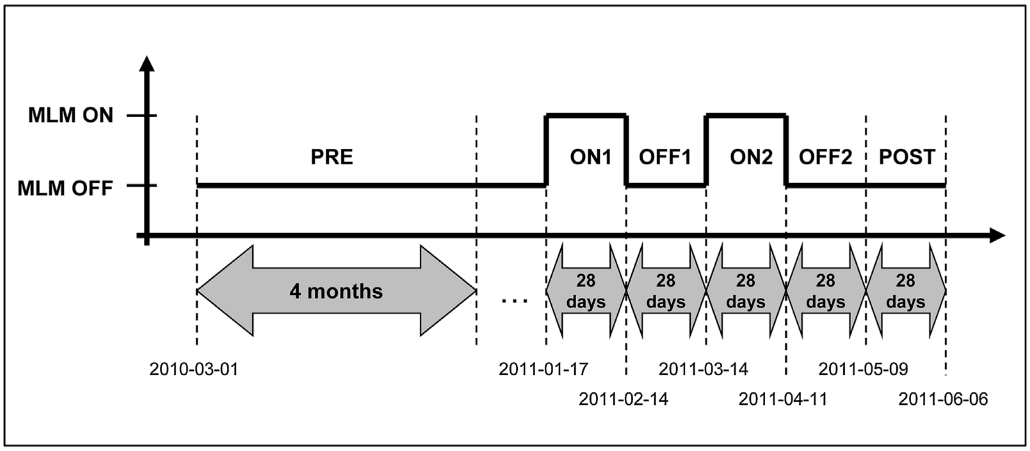


Fig. 3. Sequence of study phases (MLM = Medical Logic Module, PRE = pre-test phase, ON1 = first MLM activation, OFF1 = first MLM deactivation, ON2 = second MLM activation, OFF2 = second MLM deactivation, POST = post-test phase).

be caused by raised awareness for PCT overutilization. PRE was followed by twice activating and deactivating the PCT MLM (ON1, OFF1, ON2, OFF2) and a terminating POST phase, each of them lasting 28 days, starting on a Monday at midnight. During each ON phase the PCT MLM displayed recommendations for which patients PCT tests should be ordered. During each OFF phase and POST the PCT MLM displayed a message that decision support was currently not available.

3.5. Outcome measures

During all study phases we measured the daily number of performed PCT tests and calculated the daily PCT order rate as the number of performed PCT tests divided by the number of current ICU patients (i.e. number of occupied beds). We collected the patient census, APACHE II and SAPS II scores and daily admissions as control variables which correspond to the clinical workload that could influence physician's ordering behavior. Additionally, we analyzed adherence by calculating the percentage of observed recommendations not to measure.

3.6. Methods for data acquisition

PCT tests, patient census, admissions, SAPS II and APACHE II scores were retrieved from the PDMS database. PCT lab results are imported into the PDMS and thus digitally available. Any invocation of the PCT MLM was logged in the Arden server environment. At the end of the study we performed telephone interviews with the participants based on a semi-structured questionnaire. This questionnaire consisted of different interview sections to examine the physicians' understanding for the motivation of the PCT study, how the physician completes the PCT order list, reasons for the decreasing usage of the PCT MLM and how the PCT-recommendations could lead to the desired success.

3.7. Methods for data analysis

The distribution of PCT rates in PRE was examined by drawing a normal QQ-Plot as well as by using the Shapiro–Wilk-test to confirm compatibility with a normal distribution. Sample size estimation was based on an expected PCT rate reduction of 15%. Differences in study phases were tested for significance using the *T*-test (two sample Welch test). The null hypothesis was that during the particular study phase the PCT rate was equal to the PRE phase (alpha = 0.05, power = 0.8). The control parameters for changing workload, i.e. daily admissions, patient census, SAPS II and admission APACHE II were examined for significant alterations using Wilcoxon rank sum test, because normal distribution could not be confirmed for these parameters.

4. Results

4.1. Demographic and other study coverage data

Table 1 displays median and interquartile range of control variables for both activations of the PCT MLM. We found no significant differences in daily admissions ($p = 0.103$), APACHE II ($p = 0.341$) and SAPS II ($p = 0.736$). However, the number of occupied beds was significantly higher in ON1 than in ON2 ($p < 0.001$), with a mean difference of 12.7%.

4.2. Unexpected events during the study

During ON2 we experienced one temporary failure of the underlying replication job which feeds patient data from the PDMS to the CDSS. Thus for one day the PCT MLM was working but its recommendations were not based on the latest PCT results. Accordingly, the availability of latest PCT values was 100% in ON1 and 95.8% in ON2.

4.3. Study findings and outcome data

4.3.1. Results of PRE

In PRE we observed 2453 nursing days during 122 days. The average census was 20.1 of 25 surgical ICU beds. 1974 PCT tests were performed, corresponding to an average daily rate of 0.807. Applying the PCT rule set to all patients in PRE showed that 22% or 434 of the performed PCT tests could have been omitted.

4.3.2. Results of study phases ON/OFF/POST

During ON1, OFF1, ON2 and OFF2 we observed 576, 564, 511 and 558 nursing days. This corresponds to 376, 410, 406 and 440 performed PCT tests. Table 2 shows detailed daily data and

Table 1
Daily admissions, occupied beds, SAPS II and APACHE II (median and interquartile range) (ON1 = first phase with activated PCT MLM, ON2 = second phase with activated PCT MLM).

	Admissions	Occupied beds	APACHE II	SAPS II
ON1	6 (5–8)	20 (19–22)	15 (11–20)	29 (22–37.5)
ON2	5 (1.75–7)	18 (17.75–19.25)	15 (12–21)	29 (23–38)

Table 2
Results of the study phases ON1, OFF1, ON2, OFF2, POST (ON1 = first MLM activation, OFF1 = first MLM deactivation, ON2 = second MLM activation, OFF2 = second MLM deactivation, POST = post-test phase, D = day, B = occupied beds, *P* = number of PCT tests, R = PCT rate, ME = mean, SD = standard deviation).

	D	1	2	3	4	5	6	7	8	9	10	11	12	13	14
ON1	B	24	24	21	22	23	21	20	19	25	24	21	20	19	17
	P	13	11	14	14	13	21	16	13	14	10	16	13	13	13
	R	0.542	0.458	0.667	0.636	0.565	1.000	0.800	0.684	0.560	0.417	0.762	0.650	0.684	0.765
OFF1	B	19	18	17	20	18	17	18	19	18	18	19	21	19	21
	P	17	15	17	19	18	14	11	14	12	15	13	8	11	16
	R	0.895	0.833	1.000	0.950	1.000	0.824	0.611	0.737	0.667	0.833	0.684	0.381	0.579	0.762
ON2	B	20	18	21	17	20	18	18	21	20	18	19	19	19	19
	P	16	18	18	20	18	11	16	11	14	14	14	18	13	11
	R	0.800	1.000	0.857	1.176	0.900	0.611	0.889	0.524	0.700	0.778	0.737	0.947	0.684	0.579
OFF2	B	19	21	19	21	20	20	19	22	20	17	21	19	19	19
	P	18	20	19	19	20	23	15	14	13	15	12	15	16	21
	R	0.947	0.952	1.000	0.905	1.000	1.150	0.789	0.636	0.650	0.882	0.571	0.789	0.842	1.105
POST	B	19	18	18	20	19	19	19	18	21	21	22	20	19	17
	P	17	17	19	15	17	14	16	17	16	12	15	12	16	17
	R	0.895	0.944	1.056	0.750	0.895	0.737	0.842	0.944	0.762	0.571	0.682	0.600	0.842	1.000

	D	15	16	17	18	19	20	21	22	23	24	25	26	27	28	ME	SD
ON1	B	19	17	19	22	18	19	20	20	20	22	21	19	21	18	20.536	2.134
	P	19	15	13	16	14	12	10	11	12	13	9	11	13	14	13.429	2.574
	R	1.000	0.882	0.684	0.727	0.778	0.632	0.500	0.550	0.600	0.591	0.429	0.579	0.619	0.778	0.662	0.149
OFF1	B	22	20	22	22	19	21	22	24	22	25	22	22	20	19	20.143	2.068
	P	17	18	18	14	13	11	14	20	14	14	10	18	15	14	14.643	2.934
	R	0.773	0.900	0.818	0.636	0.684	0.524	0.636	0.833	0.636	0.560	0.455	0.818	0.750	0.737	0.733	0.155
ON2	B	19	18	18	16	18	14	16	20	18	18	22	17	16	14	18.250	1.917
	P	9	18	17	15	17	12	14	12	13	17	11	12	14	13	14.500	2.861
	R	0.474	1.000	0.944	0.938	0.944	0.857	0.875	0.600	0.722	0.944	0.500	0.706	0.875	0.929	0.803	0.173
OFF2	B	20	19	21	21	19	20	19	21	20	23	20	20	20	19	19.929	1.184
	P	16	16	14	13	17	19	15	16	10	17	12	6	12	17	15.714	3.609
	R	0.800	0.842	0.667	0.619	0.895	0.950	0.789	0.762	0.500	0.739	0.600	0.300	0.600	0.895	0.792	0.190
POST	B	19	20	21	23	19	19	17	20	22	21	22	23	22	20	19.929	1.676
	P	19	22	21	16	15	13	16	8	15	13	14	21	17	17	15.964	2.963
	R	1.000	1.100	1.000	0.696	0.789	0.684	0.941	0.400	0.682	0.619	0.636	0.913	0.773	0.850	0.807	0.166

calculated PCT rates. The average PCT rate was 0.662 in ON1, 0.733 in OFF1, 0.803 in ON2 and 0.792 in OFF2.

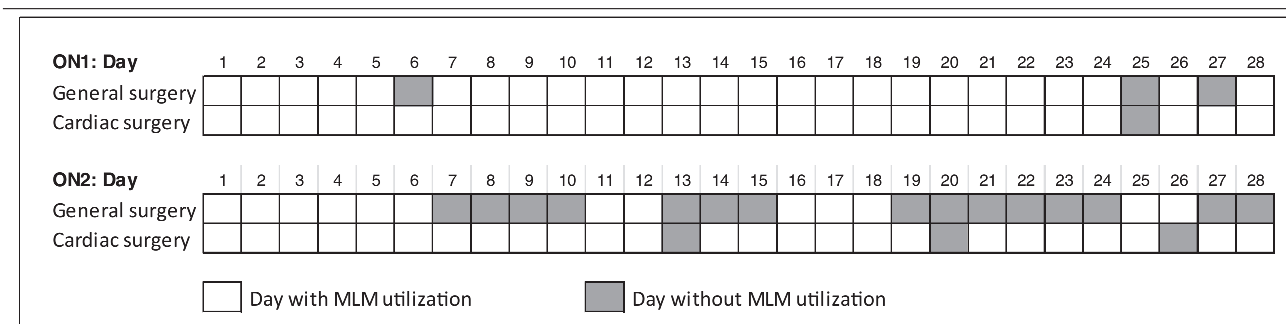
These results have been summarized in Fig. 4, which shows box plots for each study phase. During ON1 a significant reduction of PCT rates compared to PRE was observed ($p < 0.001$). This reduction continued in OFF1 ($p = 0.029$). In ON2, OFF2 and POST no significant alteration of PCT rates compared to PRE could be observed.

These results correspond to the PCT MLM utilization (Table 3). Analysis of the MLM utilization logs demonstrated that in ON1 the PCT MLM was used on all except 3 days in both sections of the ICU. In ON2, however, the utilization showed a strong decline, especially in the general surgery section with sequences of up to 6 days where the MLM was not used at all. Total usage (individual PCT tests for which the MLM was consulted) was 92% in ON1 and 62.3% in ON2.

In POST we observed 558 nursing days. 447 PCT tests were performed. The average census was 16 of 25 surgical ICU beds. The average PCT rate was 0.807, identical with PRE.

Physicians adhered to 49.5% of the recommendations in cardiac surgery, and to 31.2% in general surgery (Table 4). In total, the CDSS was invoked 90 times, 87 of these calls could be attributed to individual physicians (3 times the CDSS was invoked but not acknowledged). Five physicians used the CDSS less than 5 times. Four physicians used it up to 10 times and four used it more than 10

Table 3
Utilization of the PCT MLM in study phases ON1 and ON2 in both sections (general and cardiac surgery) of the surgical ICU (PCT = procalcitonin, ON1 = first MLM activation, ON2 = second MLM activation).



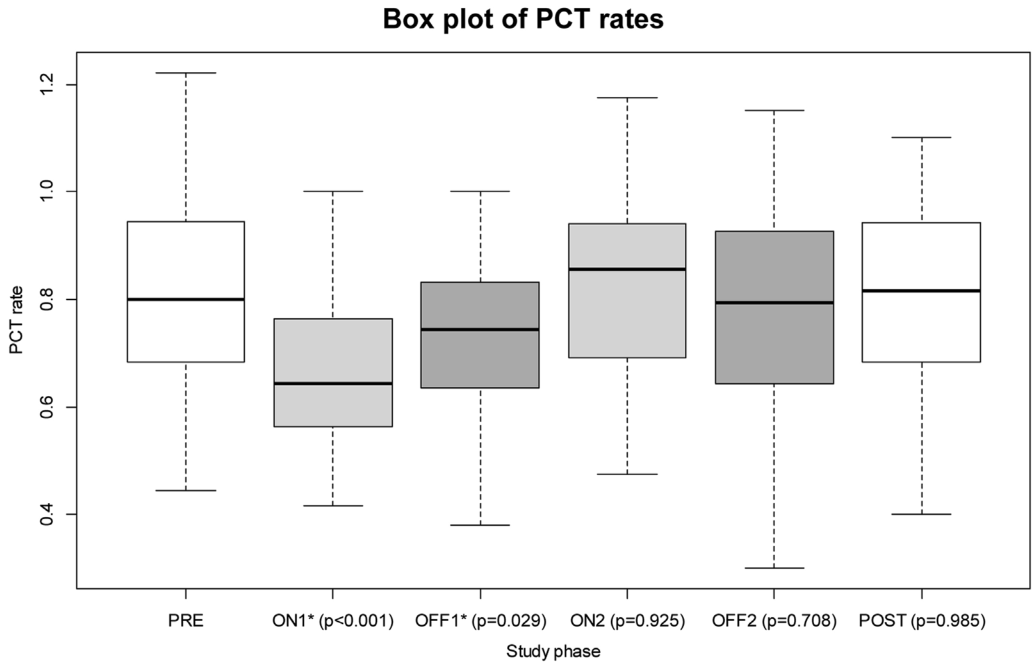


Fig. 4. PCT rates of study phases, significant values marked with * (PCT = procalcitonin, PRE = pre-test phase, ON1 = first MLM activation, OFF1 = first MLM deactivation, ON2 = second MLM activation, OFF2 = second MLM deactivation, POST = post-test phase).

times. The four main users of the CDSS had an individual compliance of 38.5%, 57.7%, 16.7% and 50%.

4.3.3. Results of Interviews

The interviews showed that only one physician knew the exact costs of a PCT test. The other physicians underestimated the costs, estimating an average cost of 6.7 euro per test, which is less than half of the actual fee. Nevertheless, the responses also demonstrated an awareness of the general problem of overutilization. Four out of six physicians emphasized that the vast majority of avoidable measurements is ordered during longer ICU stays when relatively stable patients are continuously monitored. The responses stated that in the initial days of a patient stay it is of paramount importance to closely monitor and rapidly stabilize the patients' health status. In that period it is less important if a PCT test is avoidable according to the rule set.

The interviews also showed that physicians are under strong time pressure when completing the orders list, a task which they estimated to require 8 min on average per section. One physician remarked that "there are days when you are glad about any second you save". Completion of the paper-based orders list was described as a very unpopular task. Both aspects result in a tendency towards ordering PCT simply for all patients in order to save time. Five physicians claimed to check, if possible, for the necessity of PCT orders. Two physicians stated the use of the PCT MLM to be time saving due to the list based overview that spares the need to browse through the patient records in order to examine the individual progression of PCT values. However, all physicians expressed that simply ordering PCT tests for all patients without checking for necessity is even more timesaving. All physicians expressed a considerable interest for future easier computerized order entry combined with a semi-automated checking for necessity.

Table 4
Compliance to the CDSS recommendations in cardiac and general surgery (ON1 = first phase with activated MLM, ON2 = second phase with activated MLM).

Cardiac surgery (%)	65.8	40.8	49.5
General surgery (%)	27.7	38.3	31.2

4.4. Unexpected observations

We expected PCT rates to correspond to the activation and deactivation of the PCT MLM, i.e. a decrease of PCT rates during both ON phases and an increase during both OFF phases. Instead, we found no significant decrease in ON2, but a still persisting decrease in OFF1. Unexpectedly, the rate of PCT tests in phase OFF2 was lower than the rate in phase ON2. During ON2, the median PCT rate was even higher than during PRE and POST.

In PRE we detected several cases where necessary PCT measurements were not performed. The most extreme case was one patient who went without PCT measurement for a total of 10 days. The exact number of missed PCT tests in PRE cannot be calculated, because PCT test gaps of several days' duration do not permit the conclusion which measurements would have been required as in the gaps we do not have proper input values for the rule set. We can estimate, however, that in PRE at least 54 measurements or every 38th measurement were missed, assuming that the second rule would entail a test at least every second day.

5. Discussion

5.1. Answer to study questions

The overutilization of PCT tests in the observed surgical ICU amounts to 22%. We encountered a significant reduction of PCT tests in ON1 ($p < 0.001$) during an intervention with the PCT MLM. A lasting reduction of PCT rates as observed in ON1 extrapolates to a cost savings potential of 15,000 euro per year for this SICU (based on current costs of 14 euro per test). In OFF1, although no decision support for PCT orders was available, we found a still significantly reduced PCT order rate. In a second intervention during ON2 a similar effect could not be observed again. This corresponds to the observation that the utilization of the PCT MLM decreased from 92% in ON1 to 62% in ON2. ON2, OFF2 and POST showed PCT rates which did not significantly differ from those in PRE. The absent reduction of PCT in ON2 could not be explained by an increased workload. In fact, the workload in ON1 was even slightly higher than in ON2. An accompanying interview of six physicians demonstrated a degree of unawareness of the exact PCT costs. Furthermore, it demonstrated that the physicians would not accept any additional workload for detailed checking for unnecessary lab tests.

5.2. Strengths and weaknesses of the study

The selected ON–OFF design is inferior compared to randomized controlled studies; therefore we cannot ensure that the effects measured in ON1 are actually caused by the intervention. We cannot, for example, clearly separate the impacts of raised awareness of overutilization from that of our CDSS recommendations. The lasting reduction of PCT rates in OFF1 could be an indicator for raised awareness.

Examining computerized interventions in an ICU environment, however, is quite delicate because patient safety hazards must be avoided. Therefore we abstained from randomizing because there would be no similar recommendations for patients in the control group for this important diagnostic parameter. Randomizing over physicians was declined to avoid different levels of support among clinicians. This concern extended to crossover designs. Other authors, such as Bates et al. [24] have been able to establish randomized study designs in other, maybe less critical and less regulated clinical environments. The selected design is superior, however, compared to a simple ON–OFF design which in our case would have delivered a positive result without the experience that over a longer time period the CDSS effects wear off. In addition, we believe that our study with its six phases is able to delimit the negative influence of periodic effects. The additional inclusion of workload parameters permitted to rule out some of these effects. A limitation of our study design is that we cannot distinguish between the initial intention of the ordering physician and the change of his decision caused by the CDSS recommendations. We are aware of one single study [24] which was effectively able to record both the physicians intention and his final decision within the study design.

5.3. Meaning and generalizability of the study

Some previous studies, starting with the pioneering one by Eisenberg et al. [25], demonstrated no effect of a CDSS on the prevention of unnecessary lab tests. In contrast, we detected a significant decrease of PCT rates in ON1 (success) and a potentially carryover effect in OFF1. Since these effects were eliminated in ON2, our experience strongly resembles that of May and colleagues [11]:

> In academic settings such as ours, it has been well demonstrated that the educational approach is short-lived, with promising effects disappearing shortly after cessation of the educational effort.

Physician interviews indicate that for the observed SICU checking for unnecessary lab tests is not standard clinical practice and is likely to be omitted in time critical situations, although the decision logic of the MLM was consented. In our opinion the following reasons contributed to the failure in ON2:

- The use of the PCT MLM was voluntary.
- The PCT MLM was not part of a complete electronic clinical pathway or order entry system and had to be invoked manually.
- Instructions how to start and where to find the MLM took place only once, no repetitive educational effort was made to increase attention to the problem of overutilization.
- We had justified concerns regarding overalerting and therefore did not force physicians to react to MLM recommendations. We did not even force them to use the MLM at all. This may have resulted in underalerting.

What are the requirements for a lasting positive effect of the PCT MLM? We consider the following factors essential:

- Avoidance of paper-based ordering as an obstacle to electronic decision support.
- Combination of decision support and electronic order entry.
- Automated and mandatory use of the CDSS.

Moreover, we may require additional and repeated information and education of medical staff. Therefore we consider this study as foundation for further investigation of appropriate laboratory testing at our institution. We presume that our results may be transferable to other time-critical clinical environments which have a tendency to rather perform an extra check-up than to try and save money.

Our results regarding omitted PCT tests and interview responses do not conform to those of Zhi et al. [26], who concluded that overutilization of lab tests occurs mainly during the initial phase of a patient stay (43% preventable tests) whereas only 7.4% of the lab tests in the longer stay were found to be unnecessary. In our setting, physicians indicated in the interviews that initial PCT tests for critical new ICU patients are rarely deemed unnecessary, whereas later PCT tests for stabilized patients may be considerably reduced.

Another interesting aspect compared to Zhi et al. [26] is that in our setting underutilization of PCT tests was less prominent. In Ref. [26] underutilization occurred in 44.8% while overutilization occurred in 20.6%, suggesting that patient safety may be heavily compromised by underutilization. We cannot confirm these results for our local setting. Only every 38th required PCT order was not performed, while more than every 5th PCT order could have been omitted according to our conservative rule set.

In a systematic review, Kobewka et al. [27] analyzed the influence of different interventions on laboratory test overutilization. Depending on the kind of intervention, they found substantial differences in the resulting relative reductions. For an exclusive CDSS intervention, they reported a median reduction of 22.8%, which is quite similar to our observation in ON1. There was a remarkable difference between studies in which the targeted physicians were involved in creating and implementing the intervention and those in which the physicians were not involved. The median reduction in both groups was similar, but in the group that was not involved, the reduction ranged from −27.7% to 99.7%, indicating the possibility that such an intervention may even be boycotted by physicians.

Based on our experiences, we derive the following recommendations for future studies:

- Paper-based ordering should be avoided.
- CDSS invocation should be automated to avoid additional effort.
- The number of active decisions a user has to make should be reduced to the minimum.
- The initial intention of the physician before reading the CDSS recommendation should be recorded.
- Repetitive educational efforts should be made to increase the awareness to the problem of over- and underutilization.

5.4. Unanswered and new questions

In our setting we encountered both under- and overutilization of PCT orders. Underutilization is highly critical, threatens patient safety and must therefore be completely avoided. Overutilization wastes money and should be preferably reduced. This is contrasted by physicians who deem any extra time effort for checking necessary lab tests as a problem due to heavy workload. Therefore we plan a careful simulation study for a fully automated environment where PCT lab test orders are computer-generated. To render this patient safe, i.e. to avoid underutilization, a consulting physician will oversee the procedure and check each order individually during the study period to avoid any patient risks.

The use of Arden Syntax to reduce lab overutilization seems rarely described, as a search on PubMed for the terms "arden syntax overutilization", "arden syntax cpoe" and "arden syntax order entry" yielded only one true result from 1996 [28]. Our study appears to be the first one evaluating an MLM for the reduction of overutilization. We experienced that the rule set required for our use case could be easily implemented in Arden Syntax, due to two characteristics of this language; the first one being the proximity to natural language, the second one being the provision of a set of operators tailored to the requirements of the medical domain. Thus, the rule set may be expressed by readable but concise conditional statements such as

- IF (LATEST pct) IS GREATER THAN valuelimit THEN. . .
- IF (LATEST pct) OCCURRED BEFORE 36 HOURS AGO THEN. . .
- IF SLOPE OF (LATEST 2 FROM pct) IS GREATER THAN trendlimit THEN. . .

The SLOPE operator within the third rule performs a least square regression with the results expressed as units per day. As a variety of decision support functions for CPOE may utilize rule sets of similar functionality and complexity, integrating Arden Syntax support into order entry systems may be a promising enhancement to enable physicians to add meaningful decision logic.

6. Conclusion

Our study was a success insofar as it demonstrated that an MLM with simple decision logic significantly reduced the problem of PCT overutilization at our institution. The potential maximum relative reduction would have been 22%; the observed reduction in ON1 was 18%. Our study was a partial failure insofar that this reduction did not continue during the following study phases. We may have to proceed to automate the lab order steps in order to achieve long lasting effects regarding under- or overutilization of PCT orders. In such a case, however, we must ensure that patient treatment remains under physician auspices and that we do not accidentally introduce new patient hazards by automating ICU activities. This implies that there are clear indications which orders have been automatically activated and safety fall back mechanisms to permit manual intervention at any time in the process. We hope that this will lead to a more conscious and effective use of a valuable resource.

Ethics committee approval

The Research Ethics Committee of the Faculty of Medicine, Friedrich-Alexander University Erlangen-Nuremberg, approved that there are no ethical or legal concerns over this study.

Acknowledgements

The Erlangen-Nuremberg University Chair of Medical Informatics received a grant from Dräger Medical AG & Co. KG (PDMS provider) for improving some PDMS functions in a development partnership. Medexter Healthcare GmbH has provided free support until 31.12.2014 for applying the Medexter Arden Engine for teaching and research projects at Erlangen University Hospital.

References

[1] Angus DC, van der Poll T. Severe sepsis and septic shock. N Engl J Med 2013;369:2063.
[2] Engel C, Brunkhorst FM, Bone H, Brunkhorst R, Gerlach H, Grond S, et al. Epidemiology of sepsis in Germany: results from a national prospective multicenter study. Intensive Care Med 2007;33:606–18.
[3] Reinhart K, Brunkhorst FM, Bone H, Gerlach H, Gründling M, Kreymann G, et al. Diagnose und Therapie der Sepsis. Intensivmed Notfallmed 2006;43:369–84.
[4] Fauci AS, Braunwald E, Kasper DL, Hauser SL, Longo DL, Jameson JL, et al. Harrison's principles of internal medicine. 17 ed. New York, NY: McGraw-Hill Medical; 2008.
[5] Kumar A, Roberts D, Wood KE, Light B, Parrillo JE, Sharma S, et al. Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. Crit Care Med 2006;34:1589–96.
[6] Karzai W, Oberhoffer M, Meier-Hellmann A, Reinhart K. Procalcitonin—a new indicator of the systemic response to severe infections. Infection 1997;25:329–34.
[7] BalcI C, Sungurtekin H, Gürses E, Sungurtekin U, Kaptanoglu B. Usefulness of procalcitonin for diagnosis of sepsis in the intensive care unit. Crit Care 2003;7:85–90.
[8] Nijsten MW, Olinga P, The TH, de Vries EG, Koops HS, Groothuis GM, et al. Procalcitonin behaves as a fast responding acute phase protein in vivo and in vitro. Crit Care Med 2000;28:458–61.
[9] Reinhart K, Meisner M. Biomarkers in the critically ill patient: procalcitonin. Crit Care Clin 2011;27:253–63.
[10] Catrou PG. Is that lab test necessary? Am J Clin Pathol 2006;126:335–6.
[11] May TA, Clancy M, Critchfield J, Ebeling F, Enriquez A, Gallagher C, et al. Reducing unnecessary inpatient laboratory testing in a teaching hospital. Am J Clin Pathol 2006;126:200–6.
[12] Janssens PM. Managing the demand for laboratory testing: options and opportunities. Clin Chim Acta 2010;411:1596–602.
[13] Roshanov PS, Misra S, Gerstein HC, Garg AX, Sebaldt RJ, Mackay JA, et al. Computerized clinical decision support systems for chronic disease management: a decision-maker-researcher partnership systematic review. Implement Sci 2011;6:92.
[14] Bürkle T, Castellanos I, Tech H, Prokosch HU. Implementation of a patient data management system - an evaluation study of workflow alterations. Stud HealthTechnol Inf 2010;160(Pt 2):1256–60.
[15] Castellanos I, Bürkle T, Prokosch HU, Schüttler J. Concept for the hospital-wide implementation of a patient data management system at a large clinical center-an interdisciplinary challenge. Anästh Intensivmed 2009;50:618–29.
[16] Hripcsak G, Ludemann P, Pryor TA, Wigertz OB, Clayton PD. Rationale for the Arden Syntax. Comput Biomed Res 1994;27:291–324.
[17] Hripcsak G, Clayton PD, Jenders RA, Cimino JJ, Johnson SB. Design of a clinical event monitor. Comput Biomed Res 1996;29:194–221.
[18] Hripcsak G, Cimino JJ, Johnson SB, Clayton PD. The Columbia–presbyterian medical center decision-support system as a model for implementing the Arden Syntax. Proc Annu Symp Comput Appl Med Care 1991:248–52.
[19] Ahlfeldt H, Johansson B, Linnarsson R, Wigertz O. Experiences from the use of data-driven decision support in different environments. Comput Biol Med 1994;24:397–404.
[20] Samwald M, Fehre K, de Bruin J, Adlassnig KP. The Arden Syntax standard for clinical decision support: experiences and directions. J Biomed Informatics 2012;45:711–8.
[21] Kraus S, Castellanos I, Toddenroth D, Prokosch HU, Bürkle T. Integrating Arden-Syntax-based clinical decision support with extended presentation formats into a commercial patient data management system. J Clin Monit Comput 2014;28:465–73.
[22] Wright A, Sittig DF. A four-phase model of the evolution of clinical decision support architectures. Int J Med Informatics 2008;77:641–9.
[23] Jenders RA, Shah A. Challenges in using the Arden Syntax for computer-based nosocomial infection surveillance. Proc AMIA Sym 2001:p289–293.
[24] Bates DW, Kuperman GJ, Rittenberg E, Teich JM, Fiskio J, Ma'luf N, et al. A randomized trial of a computer-based intervention to reduce utilization of redundant laboratory tests. Am J Med 1999;106:144–50.
[25] Eisenberg JM, Williams SV, Garner L, Viale R, Smits H. Computer-based audit to detect and correct overutilization of laboratory tests. Med Care 1977;15:915–21.
[26] Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaout R. The landscape of inappropriate laboratory testing: a 15-year meta-analysis. PLoS ONE 2013;8:e78962.
[27] Kobewka DM, Ronksley PE, McKay JA, Forster AJ, van Walraven C. Influence of educational, audit and feedback, system based, and incentive and penalty interventions to reduce laboratory test utilization: a systematic review. Clin Chem Lab Med 2015;53:157–83.
[28] Broverman CA, Clyman JI, Schlesinger JM, Want E. Clinical decision support for physician order-entry: design challenges. Proc AMIA Annu Fall Symp 1996:572–6.