BMJ Open

CADDIE2 - Evaluation of a clinical decision-support system for early detection of SIRS in pediatric intensive care: study protocol for a diagnostic study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-028953
Article Type:	Protocol
Date Submitted by the Author:	04-Jan-2019
Complete List of Authors:	Wulff, Antje; Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School Montag, Sara; Department of Pediatric Cardiology and Intensive Care Medicine, Hannover Medical School Steiner, Bianca; Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School Marschollek, Michael; Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School Beerbaum, Philipp; Department of Pediatric Cardiology and Intensive Care Medicine, Hannover Medical School Karch, André; Institute of Epidemiology and Social Medicine, University of Muenster Jack, Thomas; Department of Pediatric Cardiology and Intensive Care Medicine, Hannover Medical School
Keywords:	Clinical Decision Support Systems, Clinical Trial, Pediatric Intensive Care Units, Systemic Inflammatory Response Syndrome

CADDIE2 - Evaluation of a clinical decision-support system for early detection of SIRS in pediatric intensive care: study protocol for a diagnostic study

Antje Wulff[a,1], Sara Montag[b], Bianca Steiner[a], Michael Marschollek[a], Philipp Beerbaum[b], André Karch[c], Thomas Jack[b]

[a]*Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School, Carl-Neuberg-Str. 1, 30625 Hannover, Germany*

[b]*Department of Pediatric Cardiology and Intensive Care Medicine, Hannover Medical School, Carl-Neuberg-Str. 1, 30625 Hannover, Germany*

[c]*Institute of Epidemiology and Social Medicine, University of Muenster, Albert-Scheitzer-Campus-1, 48149 Muenster, Germany*

[1]Corresponding author at:
Peter L. Reichertz Institute for Medical Informatics of
TU Braunschweig and Hannover Medical School,
Carl-Neuberg-Str. 1, 30625 Hannover, Germany.
E-mail address: antje.wulff@plri.de (A.Wulff).
URL: http://www.plri.de (A. Wulff).

Email addresses for all authors:
antje.wulff@plri.de
montag.sara@mh-hannover.de
bianca.steiner@plri.de
michael.marschollek@plri.de
beerbaum.philipp@mh-hannover.de
andre.karch@ukmuenster.de
jack.thomas@mh-hannover.de

Word Count (excluding title page, abstract, references, figures and tables and captions):

3976

ABSTRACT

Introduction: Systemic inflammatory response syndrome (SIRS) is one of the most critical indicators determining the clinical outcome of pediatric intensive care patients. Clinical decision-support systems (CDSS) can be designed to support clinicians in detection and treatment. However, the use of such systems is highly discussed as they are often associated with accuracy problems and "alert fatigue". We designed a CDSS for detection of pediatric SIRS and hypothesize that a high diagnostic accuracy together with an adequate alerting will accelerate the use. Our study will (1) determine the diagnostic accuracy of the CDSS compared to "gold standard" decisions created by two, blinded experienced pediatricians, and (2) compare the system's diagnostic accuracy with that of routine clinical care decisions compared to the same gold standard.

Methods and analysis: CADDIE2 is a single-arm, controlled, prospective diagnostic accuracy study taking place at the Department of Pediatric Cardiology and Intensive Care Medicine at the Hannover Medical School and represents the second step towards our vision of cross-institutional and data-driven decision support for intensive care environments (CADDIE). The study comprises (1) recruitment of up to 300 patients, (2) creation of "gold standard" decisions, (3) routine SIRS assessments by physicians, (4) SIRS assessments by a CDSS, and (5) statistical analysis with a modified approach for determining sensitivity and specificity and comparing the accuracy results of the different diagnostic approaches.

Ethics and dissemination: Ethics approval was obtained at the study center. Results of the main trial will be communicated via publication in a peer-reviewed journal.

Discussion: We present a study design for evaluating the diagnostic accuracy of a CDSS from a routine clinical care perspective. CADDIE2 recruitment has been started successfully. In case of positive study results, our evaluation will demonstrate the potentials of CDSS use and foster the acceptance of such systems in routine decision-making.

Trial registration: ClinicalTrials.gov NCT03661450. Registered 07 September 2018. Recruitment started August 1st, 2018, and it is expected to continue until February 2019.

Protocol version: V1.0, 2018-Mar-06 Original

Keywords: Clinical Decision Support Systems, Clinical Trial, Pediatric Intensive Care Units, Systemic Inflammatory Response Syndrome

Abbreviations: AMG, "Arzneimittelgesetz" (German), Medicinal Products Act; CADDIE, Cross-institutional and Data-driven Decision Support for Intensive Care Environments; CDSS, Clinical Decision-Support System; GCP, Good Clinical Practice; GMDS, German Association for Medical Informatics, Biometry and Epidemiology; ICD, International Statistical Classification of Diseases and Related Health Problems; IPSCC, International Pediatric Sepsis Consensus Conference; PICU, Pediatric Intensive Care Unit; PDMS, Patient Data Management System; RCT, Randomized Controlled Trials; SIRS, Systemic Inflammatory Response Syndrome; STARD, Standards for Reporting Diagnostic Accuracy Studies; SPIRIT, Standard Protocol Items: Recommendations for Interventional Trials.

ARTICLE SUMMARY

Strengths and limitations of this study

- Related studies reached successful results in the context of decision-support for SIRS detection, but due to the study design, the reported results often do not reflect the usefulness of such systems in routine clinical care.
- We present an adjusted and novel approach for the course and the statistical analysis of diagnostic studies from a more routine clinical care perspective, because our study comprises
- (1) the validation of the clinical decision-support system in the comparison to the assessment of two experienced clinicians by blinded chart review ("gold standard") and
- (2) the comparison of the system's diagnostic accuracy with the diagnostic accuracy of assessments by clinicians working in routine clinical care and manually evaluating patients.
- Although our study does not comprise specific evaluations of CDSS user acceptance, it is predestinated to present the potentials of CDSS use in routine clinical care and, thus, to foster the willingness to trust the system in future.

INTRODUCTION

The first definition of systemic inflammatory response syndrome (SIRS) and sepsis in adult patients was made by the members of the "American College of Chest Physicians/Society of Critical Care Medicine Consensus Conference Committee" in 1992. SIRS was described as a "systemic inflammatory response to a variety of severe clinical insults", sepsis on the other hand was "the systemic response to infection" [1]. The criteria have been adapted to pediatric patients by the International Pediatric Sepsis Consensus Conference (IPSCC) in 2005 [2]. According to this, SIRS was present if the patient presented two or more of the defined criteria (at least an abnormal core temperature or leukocyte count are mandatory). The other criteria include abnormal heart rate, respiratory rate and blood pressure. All these criteria are specified depending on the age of the patient. According to the IPSCC, sepsis is defined as SIRS in the presence of or as a result of suspected or proven infection. Although the definition of SIRS is no longer taken into account in adult medicine, it is still relevant in pediatric medicine. Pediatric patients with SIRS and sepsis are known to have a higher risk of morbidity and mortality [3–6]. SIRS in pediatric patients also causes a significantly prolonged stay in intensive care after cardiothoracic surgery [7] and entails an increased probability of single and even multiple organdysfunction [8].

Independent of the underlying disease and together with SIRS and sepsis, organ dysfunction and failure represent the critical determinants of patient outcome in both adults and pediatric intensive care medicine. Prevention and rapid effective treatment of multi organ dysfunction and failure is crucial for survival. An optimization of the diagnostic and therapeutic workflow is very likely to have an immense impact on clinical outcome of any critically ill patient. In pediatric septic shock patients, every hour without appropriate treatment was associated with an increased risk of death by 40% [9]. Conclusively, early recognition, evaluation and treatment of pediatric SIRS and sepsis are vital for improved survival [10].

To assure that patients are treated with the best available approaches, evidence-based medicine [11] in combination with personal expertise represent the current "gold standard" of medical patient management. However, clinicians are often confronted with a stressful environment, which fosters decision-making with a lower quality than aspired [12]. This is particularly true for intensive care settings, in which clinicians are in need to make the majority of their decisions under challenging conditions characterized by an high degree of dynamics, uncertainty and risk, a need for immediate decisions and a vast amount of data.

Altogether, these factors carry risk for medical errors and adverse effects on patient safety [13–15].

To tackle these challenges, clinicians can be supported by clinical decision-support systems (CDSS) as these systems analyze, summarize and present crucial information. The growing digitalization of medical processes and patient care involves an immense amount of highly heterogeneous datasets carrying the potential to be valuable for other purposes than initially expected (*secondary use of data*): the design of systems that are able to efficiently reuse, analyze and present routine data, and thereby making data meaningful for clinical care, is fostered. CDSS are shining examples for systems processing clinical and non-clinical data and delivering an added value by detecting diseases, recommending therapies or uncovering yet unknown disease patterns [16]. Particularly in sophisticated intensive care settings, the use of highly developed automated patient data management systems (PDMS) allow the continuous recording of multiple clinical parameters and make high quality data available for the secondary use of data in CDSS.

In our previous work, we designed a CDSS for the detection of SIRS in pediatric intensive care [17]. However, only when used in routine clinical care, the potential and benefits can be fully reached and translated into clinical care. Consequently, an extensive evaluation of the system's diagnostic accuracy is needed to assure that users trust and use the system. This need is even aggravated in our context, because we strive for (1) using an automatic SIRS-labelling to train machine learning algorithms, and (2) reaching a self-learning system able to continuously process data and optimize the recommendations when used in routine care (*Learning Healthcare System*) [18]. In our previous work, we describe this approach for CDSS design, in which we denote the conduction of such a trial as **second** important step towards the vision of **c**ross-institutional **a**nd **d**ata-driven **d**ecision support for **i**ntensive care **e**nvironments (CADDIE, CADDIE2 for the presented trial) [18].

Related studies in the field of CDSS for SIRS and sepsis already reached satisfying results [19–21]. However, due to the study designs, the reported results might not reflect the usefulness of the system in routine clinical care. Often, the coding of diagnoses as ICD (International Statistical Classification of Diseases and Related Health Problems) is used as "gold standard" against what the system's results are compared to. However, in ICD, episodes of SIRS and sepsis are not documented detailed enough, e.g. the time of occurrence as well as the clinical explanations are not described. Even though sometimes additional scores are used, not all relevant episodes, which occurred during the intensive stay of a patient, can be

reflected. Hence, systems evaluated with such an approach in fact have been successfully trained, but with respects to the ICD documentation and not routine decision-making. Additionally, in many studies, the study population is preselected and does only include patients who already have the diagnosis of sepsis on admission or provide a complete data set of all vital signs and documentations, which are required as input for the algorithm used. Such perfect data sets are often not available in routine clinical care settings.

The exploration of factors influencing a successful implementation of CDSS is an ubiquitous topic. In a recent literature review, Kilsdonk et al. [22] identified possible factors influencing an successful implementation of guideline-based CDSS. One of the aspects reported mostly, deals with the *information quality of the system* and covers the relevance of data and messages delivered by the system [22]. This finding relates to a well-known and obviously still unsolved issue called "alert fatigue" [23]. Other recent work published by Liberati et al. [24] describe the conduction of a qualitative study to identify different clusters of attitudes and barriers towards CDSS implementation. The authors describe that, together with a poor integration in the clinical workflow, the "fear of experiencing excessive number of alerts" [24] is one of the factors hindering the willingness to trust systems and to believe in their unforeseen opportunities (*mutual adjustment*). This step is declared as one of the most challenging obstacles in CDSS adoption. It is suggested to integrate and evaluate the CDSS in routine clinical care and in cooperation with real users to overcome these limitations [24].

Against the background of our CADDIE objectives and the reported findings on successful CDSS implementations, we conclude that there is a need for a CADDIE2 trial focusing on validating the CDSS for SIRS and sepsis detection in routine clinical care.

Study objectives and diagnostic approaches

The **primary goal** of our study is the evaluation of the diagnostic accuracy of the CDSS for detecting SIRS in pediatric intensive care patients (*diagnostic approach I*), in comparison to the assessments of two clinicians by blinded chart review. In case of disagreement, a third clinician will be consulted. The expert assessments will be treated as "gold standard" and comprise retrospective, extensive reviews and analyses of the patient's data, state and documentations.

The **secondary goal** of our study is to compare the diagnostic accuracy of the CDSS for detecting SIRS in pediatric intensive care patients evaluated against this gold standard, to the

diagnostic accuracy of routine assessments of different clinicians working in routine clinical care (*diagnostic approach II*) when compared to the same gold standard.

Trial design and study setting

The CADDIE2 trial is designed as a single-arm, controlled, prospective diagnostic accuracy study. Single study center is the Department of Pediatric Cardiology and Intensive Care Medicine at the Hannover Medical School (monocentric). The estimated study duration is one year. Our study does not contain comparisons between different patient populations (single-arm) or interventions (no randomization). Each patient will be assessed by both diagnostic approaches. Our study can be classified as a non-drug-study (according to the German "Arzneimittelgesetz", AMG; *Medicinal Products Act*) as it does not include interventions with medicinal products.

METHODS AND ANALYSIS

Preceding studies

For our study design, we can revert to a randomized controlled trial (RCT) with 807 pediatric intensive care patients from the same ward as for the planned study. As a result of this RCT, the expected SIRS prevalence on admission to PICU was reported as 5/100 and 20-10% of the patients developed SIRS later on during their PICU stay [25]. Furthermore, we conducted a proof-of-concept study focusing on the technical practicability [17]. The proof-of-concept study yielded promising results for both the technical infrastructure and the accuracy of the system (sensitivity of 1.00, specificity of 0.94) [17].

Recommendations and guidelines

We reviewed ground work on study planning, national recommendations and templates of ethics committees and associations, and followed the *Good Clinical Practice* (GCP) in non-drug trials. We designed our study in accordance with the "Standards for Reporting Diagnostic Accuracy Studies" (STARD) [26] and the "Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) [27] guidelines (see Figure 1 in section *Timeline*, see additional file 1 for SPIRIT checklist).

Patient population and eligibility criteria

All pediatric patients aged between 0 to 18 years admitted to the study center - independent of the gender, the underlying disease or the time of admission - will be asked for their consent to participate. Patients will be recruited continuously and included, if a positive consent is available, and their length of stay exceeds twelve hours.

The physicians who will evaluate the patients during routine care are mainly specialized pediatricians with experience in pediatric intensive care. There are always one more experienced physician (working in this PICU for over a year) and one less experienced physician (working in this PICU for less than a year) in charge. The reviewers who will perform the manual chart review for creating "gold standard" decisions are specialized pediatricians and very experienced in pediatric intensive care (working in this PICU for over three years).

Outcome measures

Sensitivity and specificity *on the level of patients* will be used as primary outcome measures. As second outcome measure, sensitivity and specificity also can be determined *on the level of intensive care days*.

Statistical analysis and Sample size calculation

For the primary outcome measure, sensitivity and specificity will be determined together with *Wald confidence intervals*. The comparison will be carried out by comparing the lower bound of the confidence interval with the null hypothesis. For the secondary outcome measure, sensitivity and specificity will be determined together with confidence intervals based on *general estimating equations*. Additionally, for the secondary goal of comparing the results to routine decisions, sensitivity and specificity will be compared by means of *McNemar tests* and confidence intervals constructed based on *general estimating equations*.

For analyzing the primary outcome measure, the assessment is carried out on the patient level. This is the most conservative approach for estimating the diagnostic accuracy. The entire period of stay is considered and information are aggregated on the patient level. This leads to situations, which cannot be represented in only one cell of a contingency table. The classical four cases are amended by a new case, which occurs because the CDSS should not only

correctly assess the occurrence of a SIRS event in general but with a correct timing (e.g. SIRS event is identified within the correct shift). For example, this fifth case prevents that alert firings on day 30 of the intensive care stay will be evaluated as *true positives* if the gold standard reports a SIRS episode on day 2. Here, the CDSS did not identify the SIRS episode within the correct shift. Thus, this case is used for the determination of both *false positives* (day 30) and *false negatives* (day 2). Hence, the fifth case (false positive and false negative) can be defined as follows: *the gold standard reports at least one SIRS episode, and the CDSS detect SIRS episodes but (at least one) not within the same shift*. All other cases are defined as usual (e.g. false positive: *the gold standard reports no SIRS episode but the CDSS detects one or multiple SIRS episodes*).

Based on the different cases, the sensitivity and the specificity will be determined. For sample size calculation, the results of the proof-of-concept study were used as a basis (sensitivity 1.00 and specificity 0.94, when calculating on the level of days). For this study with the modified statistical analysis approach, a sensitivity of 90% (alternative hypothesis: 0.98, null hypothesis: 0.90) and a specificity of 80% were expected (alternative hypothesis: 0.90, null hypothesis: 0.80), with a given accuracy of estimate of 95% (type I error = 0.05) and a power of 90% (chi square test). Consequently, 97 patients suffering from at least one SIRS episode as well as 137 patients suffering from one or none SIRS episodes are required. Based on the expected incidence and prevalence, at least 300 patients need to be considered.

Timeline

Before starting the study, the clinicians were introduced in the objectives and tasks. No interventions, treatments or other care-related actions are prohibited during the trial and all patients are treated with the standard procedures (including data collection and measurements). Personal briefings on the new routine documentation were carried out during this *pilot phase* (1 month). A designated assistant physician will present the study to the patients, their parents or their legal guardians and ask for consent within the *recruiting phase*. If no consent is available, the patient will not be recruited. Simultaneously, clinicians will report their findings during their working shift per patient (routine assessments, *diagnostic approach II*). The clinicians do not perform extensive analyses of documentations or reported data (*assessment phase I*, with recruitment estimated as six months). It is documented whether the patient suffered from SIRS, sepsis or organ dysfunction (via digital documentation form,

see Figure 2). The first two weeks of this phase will be treated as test phase. If patients leave the ward after less than twelve hours, they will be excluded.

Later on, two experienced clinicians will start with their weekly, extensive, blinded review and the definition of "gold standard" assessments per patient and per shift (*assessment phase II*, at least three months). As soon as 97 patients suffering from one or more SIRS episodes as well as 137 patients suffering from one or none SIRS episodes have been identified, the recruitment will be terminated early on. Simultaneously, the data sets from all recruited patients will be integrated into a data repository to make them accessible for the CDSS. Then, the CDSS assessments per patients and per hour will start (*diagnostic approach I*). In the final *analysis phase* (at least two months), the diagnostic accuracy of the CDSS will be evaluated by comparing the assessments to the "gold standard" decisions from the experts (primary goal of the study). Additionally, the diagnostic accuracy of the routine assessments will be determined by comparing them to the same "gold standard" decisions. Afterwards, the different accuracies can be compared (secondary goal of the study). Finally, the trial results will be communicated via publication in a peer-reviewed journal. Figure 1 visualizes the different study episodes in a schematic diagram.

Figure 1: Time schedule and study episodes for CADDIE2 trial

Recruitment and consent

Patients fulfilling the eligibility criteria, their parents or their legal guardians will receive information about the study during a personal discussion with a physician. Additionally, they will get an information letter together with the consent form (available in German, English, Turkish and Arabic). Due to the pediatric specialty, we decided to create additional information sheets, one for children aged between 6 and 11, and one for children aged between 12 and 18. The families also will receive privacy statement forms (data protection, accessibility and confidentiality). An example consent form, including study information and privacy statements, can be found in additional file 2.

Patient involvement

There were no funds or time allocated for patient and public involvement so we were unable to involve patients. We intend to disseminate the results to the participants and will invite patients to help us developing an appropriate method of dissemination.

Data management and collection

The CDSS is an application with interfaces to a data repository, which is based on an semantic interoperability standard for clinical information representation (*openEHR* [28]). For more information on data processing and managing, we refer to our previously published paper [17]. For reasoning, the CDSS queries data from and stores results into the repository. For the routine assessment (assessment phase I, diagnostic approach II), we created a documentation form, which is based on the same interoperability standard and the same interfaces to the data repository (see Figure 2). Thereby, all results (gold standard, diagnostic approach I "CDSS", diagnostic approach II "routine assessments") will be available in the same interoperable format. Patient data (identification, birthdate), intensive care parameters (heart rate, respiratory rate, body temperature, leukocytes, neutrophils, mechanical ventilation, cooling devices, pacemaker), alert parameters (SIRS, sepsis, organ dysfunction with duration, beginning and end of the episode and shift), and general documentations of the patient conditions, events or unintended effects will be documented and processed.

Figure 2: Digital documentation form (based on openEHR data repository)

Data monitoring and auditing

Accompanying quality assurance measures are continuously carried out to ensure high data quality. Plausibility checks will be executed while integrating data into our repository (e.g. simple counts can uncover whether data from the primary source systems is missing in the data repository). Furthermore, the data set will be randomly reviewed by physicians to guarantee the plausibility from a clinical perspective. By following the openEHR standard, we

are able to automatically execute validation checks on specific parameters (e.g. definition of ranges for specific values, or double data entries). Thereby, missing or wrong values will be uncovered automatically when integrating the data sets or filling out the documentation form. The trial procedures will be monitored continuously by the authors as well as by designated physicians and nurses. They will supervise the adherence to the study protocol, the procedures for routine documentation and data integration, data quality and privacy.

Data protection: data access and confidentiality

We designed a data protection concept in cooperation with the local data security officer. The concept defines pseudonymization and data access procedures, outlines the patient consent, and explains technical security mechanisms. All data sets collected or created as part of this trial are treated as strictly confidential. The data sets will be stored pseudonymized and in secure conditions in the data repository located in the network of the Hannover Medical School. To prevent unauthorized disclosure of patient information, it is only accessible for the physicians and employees in charge for this study. Collected data from patients withdrawing their consent (drop outs), will be completely deleted from the data repository. All study files, the final trial data sets as well as the trial results will be archived for ten years in an approved long-term repository and in accordance to the relevant legal and statutory requirements. The patient will be informed about these procedures as well as their rights (including the possibility to withdraw the consent and to obtain information about collected data sets at any time), and will be asked to consent to these (see section Recruitment and consent).

Ethics and dissemination

All aspects are designed according to the General Data Protection Regulation from the European Union (2016/679) and are accepted by the data security officer of the Hannover Medical School. A positive vote of the ethics committee was given (No. 7804_BO_S_2018, see Additional file 3). Further details on data protection aspects can be requested from the authors. Results of the main trial will be communicated via publication in a peer-reviewed journal. Furthermore, we intend to disseminate the results to the participants through an appropriate method of dissemination to be defined.

DISCUSSION

To be used in the long run, CDSS have to deliver relevant information in a timely manner and at an adequate frequency. Current approaches for evaluating the usefulness of CDSS indeed present positive results. However, due to a restricted study design, which is not designed towards the specific conditions of daily work, evaluation results may not represent the feasibility of the system in routine clinical care. Hence, with our work, we contribute a modified study design for evaluating the diagnostic accuracy of a CDSS with a strong focus on routine clinical care. We hypothesize that such an evaluation will demonstrate the potentials of CDSS use in routine clinical care. In case of a positive study outcome, we will be able to reason that our CDSS is not only feasible from a technical but also from a clinical perspective as it supports clinicians in critical diagnostic decision-making. For evaluating the diagnostic accuracy of a CDSS during a clinical trial, a so-called "gold standard" representing the true state of the patient is required. However, in complex, knowledge-and experience-based contexts as diagnostic decision-making, reproducible, objective and quantitative "gold standards" are rare. In contrast to related trials, we use an excessive evaluation of the patient by two experience clinicians as benchmark. To reduce possible biases, the clinicians are blinded to each other as well as to the CDSS. In situations of disagreement, a third clinician will be consulted, decisions will be revealed and a consensus decision will be reported. We are aware that our approach is time-consuming requiring highly engaged clinicians. Because of the stressful intensive care environment, assessments may be delayed, and thus, the timeline of the study may not be adhered. For an early recognition of issues and study monitoring at the ward, an assistant physician is in charge. Also the routine assessments of clinicians have to be managed as they are at the same risk to be biased. Clinical documentation might be handled more meticulous at the beginning and more careless in the end of the study. With respect to the new documentation form, a positive feedback was given within the test phase of the study. This might be due to the fact that clinicians were involved in designing the form. Furthermore, the form was integrated in the PDMS used daily. Together with the designated study monitor, this integration raised the satisfactory and the utilization rate of the form as well as the adherence to the study protocol.

For sample size calculation, it might be possible that the incidence was overestimated, so that in our settings more than 300 patients are needed to reach 97 patients suffering from at least one SIRS episode. The recruitment will be continuously aligned towards the number of recruited SIRS patients to be able to stop the recruitment as soon as the required number of

SIRS patients has been reached. Our expected values for sensitivity and specificity are rather conservative because we decided to primarily use an equally conservative statistical analysis approach by calculating at the level of patients. However, the expected values are treated as acceptable in clinical routine as the diagnostic accuracy of the system will be over a critical minimum (and with respect to the aspired second goal of our study, even better than in clinical routine decision-making). At the same time, alert fatigue will be prevented because specificity is equally high. Current CDSS often indeed present a higher sensitivity but a very low specificity. We are confident that our thoughts meet the need for an optimum balance between sensitivity and specificity, e.g. as reported by Coleman et al. [23]. Nevertheless, we will enhance our results with a more liberal analysis on the level of days.

Our study has been successfully started with recruitment according to this design and promises valuable results. When reaching a good diagnostic accuracy compared to the gold standard as well as advantages over the diagnostic accuracy of routine assessments, we are optimistic that our users are willing to trust and use the system in future.

STATEMENTS

Author Contributions

AW was responsible for design and implementation of the presented clinical decision-support system and the outline of the study protocol, and has drafted the manuscript. TJ provided

clinical expertise for the use case and the design of the underlying knowledge model, leaded the proof-of-concept study and co-drafted the manuscript. AK was primarily responsible for the design of the statistical analysis, the sample size calculation and the authoring of the corresponding sections. BS and SM helped in the conception of the general study approach but especially for definition of goals and outcome measures, timing and patient recruitment; SM is responsible for patient recruitment and monitors the study at the ward. PB and MM provided clinical expertise for study design, revised the manuscript critically, and gave subject-specific advices as well as the final approval of the manuscript version to be published. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Funding

No funding to declare.

Availability of data and material

The datasets generated and/or analyzed during the current study are not publicly available due to data privacy and security matters of patients but are available from the corresponding author on reasonable request.

Ethical approval and patient consent

This diagnostic study received ethics approval from the Hannover Medical School Ethics Committee (approval number 7804_BO_S_2018). The trial is registered with ClinicalTrials.gov (NCT03661450).

Consent to participate will be given by the patients, by their parents or by their legal guardians by signing a study consent form. All aspects are designed according to the General Data Protection Regulation (EU) 2016/679 and are accepted by the data security officer of the Hannover Medical School as well as by the Ethics Committee.

Protocol modifications will require a formal amendment to the protocol which will be reported to the Hannover Medical School Ethics Committee for approval.

Acknowledgements

We want to thank our colleagues from the ZIMt department for Educational and Scientific IT Systems of the Hannover Medical School for supporting in matters of data access and integration.

REFERENCES

1 American College of Chest Physicians/Society of Critical Care Medicine Consensus Conference: definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. *Critical care medicine* 1992;20(6):864–74. PubMed PMID: 1597042.

2 Goldstein B, Giroir B, Randolph A. International pediatric sepsis consensus conference: Definitions for sepsis and organ dysfunction in pediatrics*. *Pediatric Critical Care Medicine* 2005;6(1):2–8. PubMed PMID: 15636651.

3 Wolfler A, Silvani P, Musicco M, et al. Incidence of and mortality due to sepsis, severe sepsis and septic shock in Italian Pediatric Intensive Care Units: A prospective national survey. *Intensive Care Med* 2008;34(9):1690–97. doi: 10.1007/s00134-008-1148-y.

4 Cvetkovic M, Lutman D, Ramnarayan P, et al. Timing of death in children referred for intensive care with severe sepsis: Implications for interventional studies. *Pediatric Critical Care Medicine* 2015;16(5):410–17. doi: 10.1097/PCC.0000000000000385.

5 Creedon JK, Vargas S, Asaro LA, et al. Timing of Antibiotic Administration in Pediatric Sepsis. Pediatr Emerg Care 2018. doi: 10.1097/PEC.0000000000001663.

6 Kortz TB, Sawe HR, Murray B, et al. Clinical Presentation and Outcomes among Children with Sepsis Presenting to a Public Tertiary Hospital in Tanzania. *Front Pediatr* 2017;5:278. doi: 10.3389/fped.2017.00278.

7 Boehne M, Sasse M, Karch A, et al. Systemic inflammatory response syndrome after pediatric congenital heart surgery: Incidence, risk factors, and clinical outcome. *J Card Surg* 2017;32(2):116–25. doi: 10.1111/jocs.12879.

8 Ganjoo S, Ahmad K, Qureshi UA, et al. Clinical Epidemiology of SIRS and Sepsis in Newly Admitted Children. *Indian J Pediatr* 2015;82(8):698–702. doi: 10.1007/s12098-014-1618-x.

9 Han YY, Carcillo JA, Dragotta MA, et al. Early reversal of pediatric-neonatal septic shock by community physicians is associated with improved outcome. *Pediatrics* 2003;112(4):793–99. PubMed PMID: 14523168.

10 Paul R. Recognition, Diagnostics, and Management of Pediatric Severe Sepsis and Septic Shock in the Emergency Department. *Pediatric Clinics of North America* 2018;65(6):1107–18. doi: 10.1016/j.pcl.2018.07.012.

11 Sackett DL. Evidence-based medicine. *Seminars in Perinatology* 1997;21(1):3–5. PubMed PMID: 9190027.

12 Zavala AM, Day GE, Plummer D, et al. Decision-making under pressure: Medical errors in uncertain and dynamic environments. *Aust Health Rev* 2017. doi: 10.1071/AH16088.

13 Lighthall GK, Vazquez-Guillamet C. Understanding Decision Making in Critical Care. Clin Med Res 2015;13(3-4):156–68. doi: 10.3121/cmr.2015.1289.

14 Castaneda C, Nalley K, Mannion C, et al. Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine. *J Clin Bioinforma* 2015;5:4. doi: 10.1186/s13336-015-0019-3.

15 Lincoln P, Manning M-J, Hamilton S, et al. A pediatric critical care practice group: use of expertise and evidence-based practice in identifying and establishing "best" practice. *Crit Care Nurse* 2013;33(2):85–87. doi: 10.4037/ccn2013740.

16 Shortliffe EH. Biomedical Informatics: Defining the Science and Its Role in Health Professional Education. In: Hutchison D, Kanade T, Kittler J, et al., eds. Information Quality in e-Health. Berlin, Heidelberg: Springer Berlin Heidelberg 2011:711–14.

17 Wulff A, Haarbrandt B, Tute E, et al. An interoperable clinical decision-support system for early detection of SIRS in pediatric intensive care using openEHR. *Artif Intell Med* 2018;89:10–23. doi: 10.1016/j.artmed.2018.04.012.

18 Wulff A, Marschollek M. Learning Healthcare Systems in Pediatrics: Cross-Institutional and Data-Driven Decision-Support for Intensive Care Environments (CADDIE). *Stud Health Technol Inform* 2018;251:109–12. PubMed PMID: 29968614.

19 Faisal M, Scally A, Richardson D, et al. Development and External Validation of an Automated Computer-Aided Risk Score for Predicting Sepsis in Emergency Medical Admissions Using the Patient's First Electronically Recorded Vital Signs and Blood Test Results. *Critical care medicine* 2018;46(4):612–18. doi: 10.1097/CCM.0000000000002967.

20 Austrian JS, Jamin CT, Doty GR, et al. Impact of an emergency department electronic sepsis surveillance system on patient mortality and length of stay. *J Am Med Inform Assoc* 2018;25(5):523–29. doi: 10.1093/jamia/ocx072.

21 Mao Q, Jay M, Hoffman JL, et al. Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department, general ward and ICU. *BMJ Open* 2018;8(1):e017833. doi: 10.1136/bmjopen-2017-017833.

22 Kilsdonk E, Peute LW, Jaspers MW. Factors influencing implementation success of guideline-based clinical decision support systems: A systematic review and gaps

analysis. *International Journal of Medical Informatics* 2017;98:56–64. doi: 10.1016/j.ijmedinf.2016.12.001.

23 Coleman JJ, van der Sijs H, Haefeli WE, et al. On the alert: Future priorities for alerts in clinical decision support for computerized physician order entry identified from a European workshop. *BMC Med Inform Decis Mak* 2013;13:111. doi: 10.1186/1472-6947-13-111.

24 Liberati EG, Ruggiero F, Galuppo L, et al. What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. *Implement Sci* 2017;12(1):113. doi: 10.1186/s13012-017-0644-2.

25 Jack T, Boehne M, Brent BE, et al. In-line filtration reduces severe complications and length of stay on pediatric intensive care unit: A prospective, randomized, controlled trial. *Intensive Care Med* 2012;38(6):1008–16. doi: 10.1007/s00134-012-2539-7.

26 Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: Explanation and elaboration. *BMJ Open* 2016;6(11):e012799. doi: 10.1136/bmjopen-2016-012799.

27 Chan A-W, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: Guidance for protocols of clinical trials. *BMJ* 2013;346:e7586. doi: 10.1136/bmj.e7586.

28 Thomas Beale. Archetypes: Constraint-based Domain Models for Future-proof Information Systems. In: Eleventh OOPSLA workshop on behavioral semantics: serving the customer. Seattle, Washington, Boston: Northeastern University 2002:16–32.

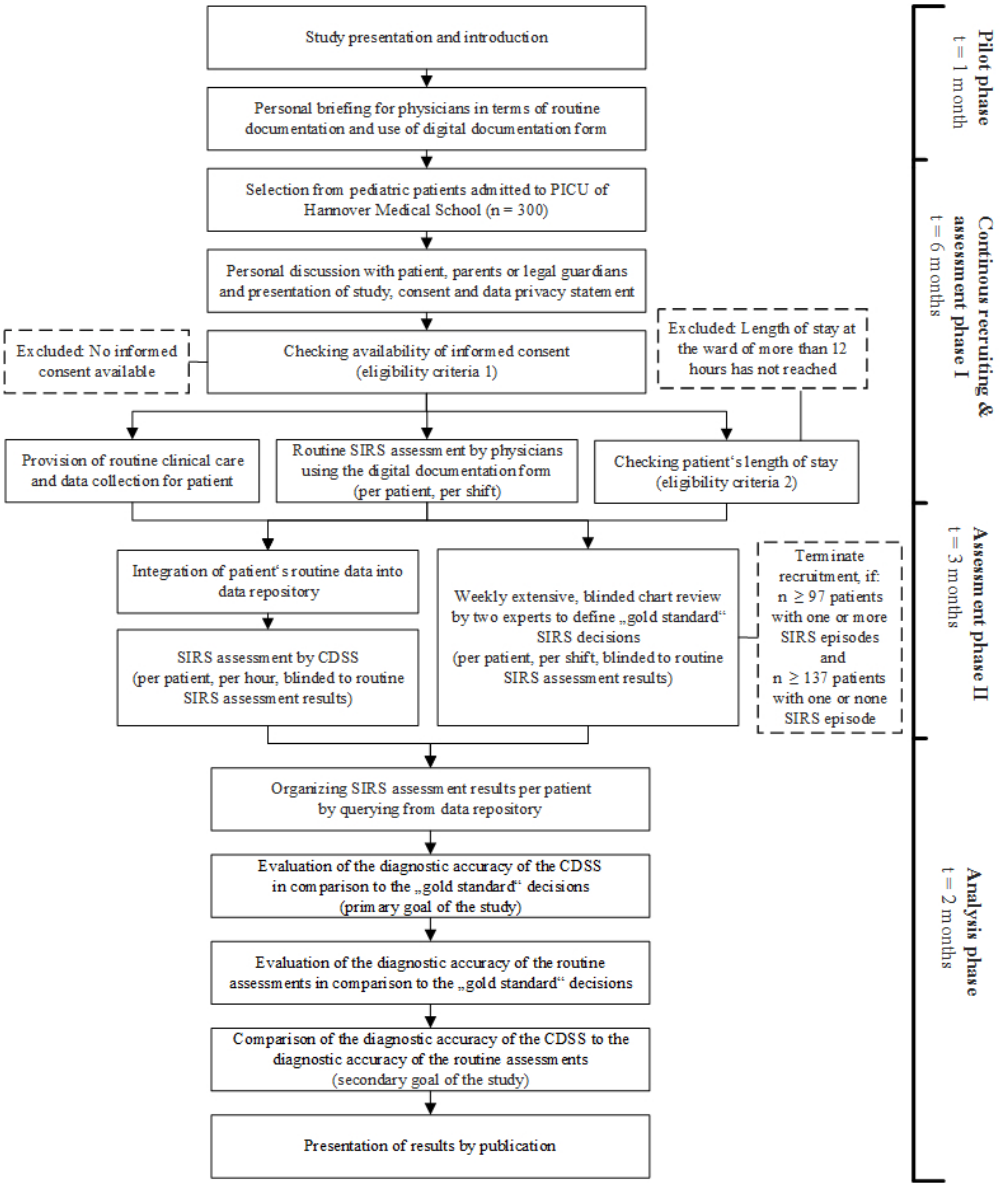

Figure 1: Time schedule and study episodes for CADDIE2 trial

208x245mm (96 x 96 DPI)

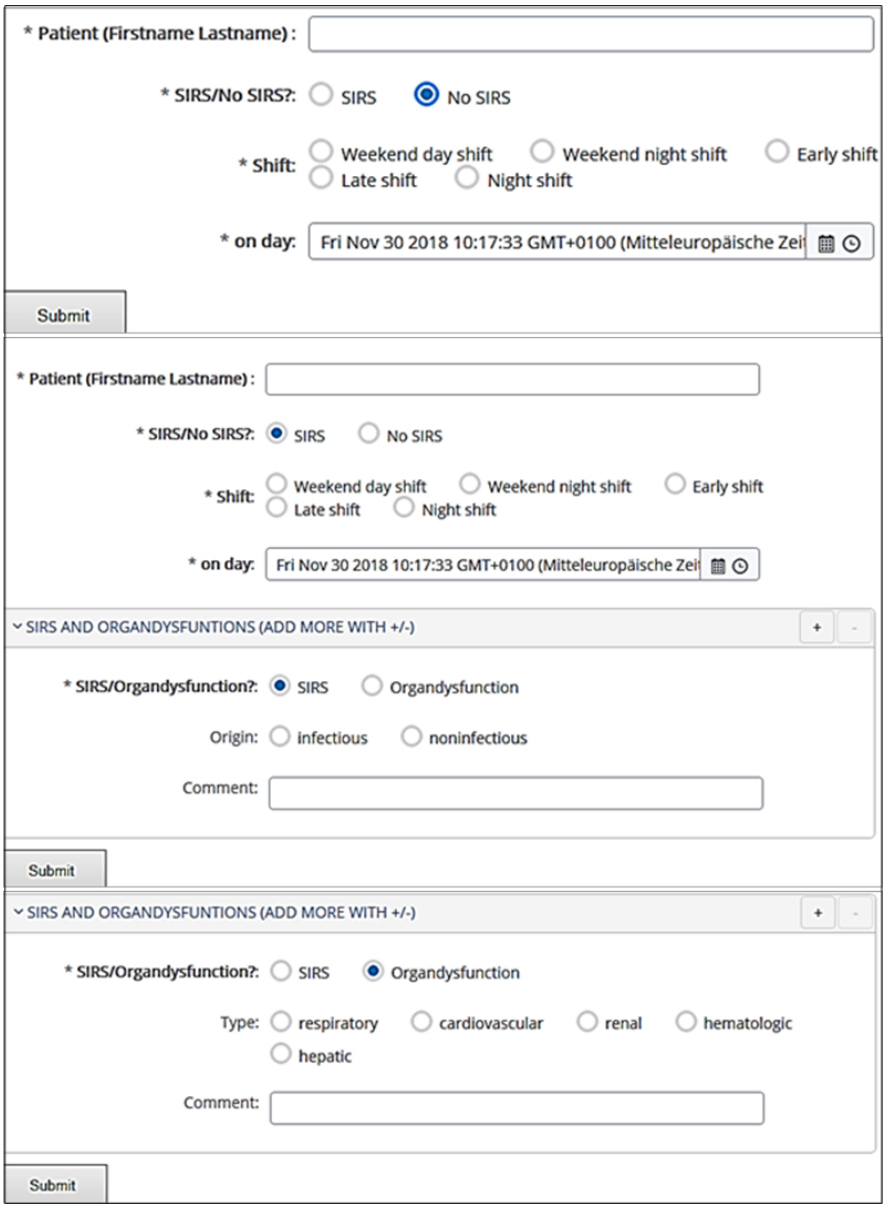

Figure 2: Digital documentation form (based on openEHR data repository)

210x286mm (96 x 96 DPI)

SPIRIT
STANDARD PROTOCOL ITEMS: RECOMMENDATIONS FOR INTERVENTIONAL TRIALS

SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*

Section/item	Item No	Description	Addressed on page number
Administrative information			
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	1
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry	3
	2b	All items from the World Health Organization Trial Registration Data Set	NA
Protocol version	3	Date and version identifier	3
Funding	4	Sources and types of financial, material, and other support	18
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	1,18
	5b	Name and contact information for the trial sponsor	NA
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	18
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	9,15,16

Introduction			
Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	_____5,6,7__
	6b	Explanation for choice of comparators	_____5,6,7,8__
Objectives	7	Specific objectives or hypotheses	_____7,8_____
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)	_____8_____
Methods: Participants, interventions, and outcomes			
Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	_____8_____
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	_____9_____
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	NA (no interventions, diagnostic study), alternative: _____8_____
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)	_____NA (each patient will be assessed by both diagnostic approaches), alternative: _____8_____

	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)	_____NA (no intervention, routine care), alternative: _____16____
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	_____10____
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	_____9,10___
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	_____10,11,12_
Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	_____9,10____
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size	_____12,13__
Methods: Assignment of interventions (for controlled trials)			
Allocation:			
Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	_____NA____
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	_____NA____
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	_____NA____

Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	7,11,16
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	16
Methods: Data collection, management, and analysis			
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	13,14
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	NA (no interventions)
Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	13,14, 15
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	9,10
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	NA
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	NA
Methods: Monitoring			
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	15
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	NA

Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	15
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	15,16
Ethics and dissemination			
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval	16, 18, 19, Additional file 3
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)	19
Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	12,13
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	NA
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	13,15
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site	18
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	18
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	NA
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	11,16

	31b	Authorship eligibility guidelines and any intended use of professional writers	_____18____
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	_____NA___
Appendices			
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates	_____Additional file 2__
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	_____NA____

*It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items. Amendments to the protocol should be tracked and dated. The SPIRIT checklist is copyrighted by the SPIRIT Group under the Creative Commons "Attribution-NonCommercial-NoDerivs 3.0 Unported" license.

Patient Information (Parents)

Dear parents,

we would like to include your child in a clinical trial, you will find the details on the following pages. Please read them carefully. If you have any questions please do not hesitate to ask the doctor in charge.

Information about the clinical trial which is named:

Evaluation of the diagnostic accuracy of a Clinical Decision-Support System (CDSS) to support recognition of SIRS and Sepsis in paediatric intensive care patients in comparison to medical specialists.

Your child was or will be admitted on the paediatric intensive care unit (PICU) at Medizinische Hochschule Hannover. Due to neccessary monitoring and therapy we will gather laboratory results and vital signs continueally.

During the clinical trial these laboratory and monitoring details will be analyzed with regard to a pronounced, generalized inflammatory reaction (Systemic inflammatory response syndrome (SIRS)) using the support of a computer-based system. The objective is to support the doctors in early recognition of SIRS.

The recognition of SIRS on a PICU is a complex task that is quite suitable for a Clinical Decision-Support System. SIRS is caused by different reasons, for example surgery, injurys or infection. It is a typical occurence at PICU-patients and puts the patient under additional strain independent from the original reason for admittance. For diagnosing SIRS in paediatric patients correctly we have to find abnormalities in different vital signs and laboratory results which have age related reference values, making it an advanced task.

The development and clinical trial of a CDSS for SIRS in paediatric patients is the result of a cooperation between the Department of Paediatric Cardiology and Intensive Care Medicine of Medizinische Hochschule Hannover and the Peter L. Reichertz Institut für

Medizinsche Informatik der Technischen Universität Braunschweig und der Medizinischen Hochschule Hannover (PLRI).

Are there any risks by taking part in the clinical trial?

The CDSS will only analyze data that is taken and saved during your child's admittance on the PICU anyway. There won't be any additional blood tests, monitoring or examinations. These standardized raised data will be saved in a pseudonymized way that ensures the patient's privacy.

What happens if I don't agree to take part in the clinical trial?

The whole team will respect your decision if you do not want your child to take part in the clinical trial. This will not be evaluated or have any negative effect on your child's treatment. You can also revoke your agreement anytime.

Privacy Policy

The data raised and saved for the trial are pseudonymized and it is impossible to reconnect them to an individual. The data will be saved for 10 years and deleted afterwards. The issues of the Bundesdatenschutzgesetz are considered in every respect.

I received the information and consent to the clinical trial mentioned above.

☐ Yes ☐ No

_____

(Date and signature of the parent/legal guardian)

_____

(Date and signature of the medical doctor)

Patient Information (Patient)

Dear patient,

we would like to include you in a clinical trial, you will find the details on the following pages. Please read them carefully. If you have any questions please do not hesitate to ask the doctor in charge.

Information about the clinical trial which is named:

Evaluation of the diagnostic accuracy of a Clinical Decision-Support System (CDSS) to support recognition of SIRS and Sepsis in paediatric intensive care patients in comparison to medical specialists.

You were or will be admitted on the paediatric intensive care unit (PICU) at Medizinische Hochschule Hannover. To treat you best we will monitor your vital signs as heartbeat, blood pressure and respiratory rate and take blood tests to monitor your white blood count and other inflammatory parameters. The results will be saved.

We know about the occurence of extreme inflammatory reactions in paediatric patients. These can be cause by several reasons as surgery, injury or infection. We can measure these by changes in your vital signs, temperature or white blood count.
With a clinical trial something new is tested, for example some treatment. For this clinical trial your monitoring or vital sign data will be analyzed by a computer programme to support your doctors to detect changes that might be caused by an inflammatory reaction as soon as possible. The sooner we are aware we can adjust your treatment.

Are there any risks by taking part in the clinical trial?

The CDSS will only analyze data that is taken and saved during your admittance on the PICU anyway. There won't be any additional blood tests, monitoring or examinations. These standardized raised data will be saved in a pseudonymized way that ensures the patient's privacy.

What happens if I don't agree to take part in the clinical trial?

The whole team will respect your decision if you do not want to take part in the clinical trial. This will not be evaluated or have any negative effect on your treatment. You can also revoke your agreement anytime.

Privacy Policy

The data raised and saved for the trial are pseudonymized and it is impossible to reconnect them to an individual. The data will be saved for 10 years and deleted afterwards. The issues of the Bundesdatenschutzgesetz are considered in every respect.

I received the information and consent to the clinical trial mentioned above.

☐ Yes ☐ No

_____

(Date and signature of the patient)

_____

(Date and signature of the medical doctor)

Medizinische Hochschule Hannover

MHH Ethikkommission OE 9515
30623 Hannover

Herrn
Dr. Thomas Jack
Klinik für Pädiatrische Kardiologie und Pädiatrische Intensivmedizin
OE 6730 – im Hause

Ethikkommission
Vorsitzender:
Prof. Dr. Stefan Engeli

Sekretariat:
Marion Lange
Telefon: 0511 532-3443
Liane Höft
Telefon: 0511 532-9812

Fax: 0511 532-16 3443
ethikkommission@mh-hannover.de

Carl-Neuberg-Straße 1
30625 Hannover
Telefon: 0511 532-0
www.mh-hannover.de

20.06.2018/MLa

Nr. 7804_BO_S_2018
Evaluation der diagnostischen Genauigkeit eines Clinical Decision-Support Systems (CDSS) zur Unterstützung der Erkennung von SIRS und Sepsis auf der pädiatrischen Intensivstation im Vergleich zu medizinischem Fachpersonal

Sehr geehrter Herr Kollege Jack,

die Mitglieder der Ethikkommission haben auf ihrer Sitzung am 21.03.2018 über o.g. Antrag beraten. Nach Eingang der gemäß unserem Schreiben vom 26.03.2018 überarbeiteten Unterlagen bestehen keine Bedenken gegenüber der Durchführung der Studie.

Die Ethikkommission weist darauf hin, dass die ärztliche und juristische Verantwortung bei den jeweiligen Prüfärzten verbleibt.

Datenschutzrechtliche Aspekte von Forschungsvorhaben werden durch die Ethikkommission grundsätzlich nur kursorisch geprüft. Dieses Votum / diese Bewertung ersetzt mithin nicht die Konsultation des zuständigen Datenschutzbeauftragten.

Mit besten Grüßen

Prof. Dr. Stefan Engeli
Vorsitzender der Ethikkommission

Folgende Mitglieder haben an der Beratung des o.g. Antrages mitgewirkt:

Prof. Dr. St. Engeli (Vorsitzender), Klinische Pharmakologie der MHH
PD Dr. U.-V. Albrecht, Stellv. Leiter des Peter L. Reichertz Institut für Medizinische Informatik
Prof. Dr. A. M. Das, Leiter der Pädiatrischen Stoffwechselmedizin, Klinik für Päd. Nieren-, Leber- und Stoffwechselerkrankungen, Zentrum Kinderheilkunde und Jugendmedizin der MHH
Dr. J. Graubner, Arzt für Allgemeinmedizin
Prof. Dr. Thomas Illig, Leiter Biobank (HUB)
Prof. Dr. A. Koch, Leiter der Abt. Biometrie der MHH
Frau Prof. Dr. B. Lohff, Em. Leiterin der Abt. Geschichte, Ethik und Philosophie der Medizin der MHH
Dr. Oliver Pramann, Fachanwalt für Medizinrecht
Jörg Viering, Leiter der Forschungswerkstätten, MHH
Prof. Dr. Peter Vogt, Leiter der Klinik für Plastische und Wiederherstellungschirurgie, MHH

BMJ Open

CADDIE2 - Evaluation of a clinical decision-support system for early detection of systemic inflammatory response syndrome in pediatric intensive care: study protocol for a diagnostic study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-028953.R1
Article Type:	Protocol
Date Submitted by the Author:	04-Apr-2019
Complete List of Authors:	Wulff, Antje; Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School Montag, Sara; Department of Pediatric Cardiology and Intensive Care Medicine, Hannover Medical School Steiner, Bianca; Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School Marschollek, Michael; Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School Beerbaum, Philipp; Department of Pediatric Cardiology and Intensive Care Medicine, Hannover Medical School Karch, André; Institute of Epidemiology and Social Medicine, University of Muenster Jack, Thomas; Department of Pediatric Cardiology and Intensive Care Medicine, Hannover Medical School
Primary Subject Heading:	Paediatrics
Secondary Subject Heading:	Health informatics, Intensive care, Infectious diseases, Diagnostics
Keywords:	Clinical Decision Support Systems, Clinical Trial, Systemic Inflammatory Response Syndrome, Paediatric intensive & critical care < INTENSIVE & CRITICAL CARE

CADDIE2 - Evaluation of a clinical decision-support system for early detection of systemic inflammatory response syndrome in pediatric intensive care: study protocol for a diagnostic study

Antje Wulff[a,1], Sara Montag[b], Bianca Steiner[a], Michael Marschollek[a], Philipp Beerbaum[b], André Karch[c], Thomas Jack[b]

[a]*Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School, Braunschweig and Hannover, Germany*

[b]*Department of Pediatric Cardiology and Intensive Care Medicine, Hannover Medical School, Hannover, Germany*

[c]*Institute of Epidemiology and Social Medicine, University of Muenster, Muenster, Germany*

[1]Corresponding author at:
Peter L. Reichertz Institute for Medical Informatics of
TU Braunschweig and Hannover Medical School,
Carl-Neuberg-Str. 1, 30625 Hannover, Germany.
E-mail address: antje.wulff@plri.de (A.Wulff).
URL: http://www.plri.de (A. Wulff).

Email addresses for all authors:
antje.wulff@plri.de
montag.sara@mh-hannover.de
bianca.steiner@plri.de
michael.marschollek@plri.de
beerbaum.philipp@mh-hannover.de
andre.karch@ukmuenster.de
jack.thomas@mh-hannover.de

Word Count (excluding title page, abstract, references, figures and tables and captions):

3995

ABSTRACT

Introduction: Systemic inflammatory response syndrome (SIRS) is one of the most critical indicators determining the clinical outcome of pediatric intensive care patients. Clinical decision-support systems (CDSS) can be designed to support clinicians in detection and treatment. However, the use of such systems is highly discussed as they are often associated with accuracy problems and "alert fatigue". We designed a CDSS for detection of pediatric SIRS and hypothesize that a high diagnostic accuracy together with an adequate alerting will accelerate the use. Our study will (1) determine the diagnostic accuracy of the CDSS compared to gold standard decisions created by two, blinded experienced pediatricians, and (2) compare the system's diagnostic accuracy with that of routine clinical care decisions compared to the same gold standard.

Methods and analysis: CADDIE2 is a prospective diagnostic accuracy study taking place at the Department of Pediatric Cardiology and Intensive Care Medicine at the Hannover Medical School; it represents the second step towards our vision of cross-institutional and data-driven decision support for intensive care environments (CADDIE). The study comprises (1) recruitment of up to 300 patients (started *August 1, 2018*), (2) creation of gold standard decisions (planned start date *February 1, 2019*), (3) routine SIRS assessments by physicians (started with recruitment), (4) SIRS assessments by a CDSS (planned start date *February 1, 2019*), and (5) statistical analysis with a modified approach for determining sensitivity and specificity and comparing the accuracy results of the different diagnostic approaches (planned start date *May 1, 2019*).

Ethics and dissemination: Ethics approval was obtained at the study center (Ethics Committee of the Hannover Medical School). Results of the main study will be communicated via publication in a peer-reviewed journal.

Trial registration: ClinicalTrials.gov NCT03661450. Registered 07 September 2018. Recruitment started August 1st, 2018, and it is expected to continue until February 2019.

Protocol version: V1.0, 2018-Mar-06 Original

Keywords: Clinical Decision Support Systems, Clinical Trial, Paediatric intensive & critical care, Systemic Inflammatory Response Syndrome

Abbreviations: CADDIE, Cross-institutional and Data-driven Decision Support for Intensive Care Environments; CDSS, Clinical Decision-Support System; GCP, Good Clinical Practice; GMDS, German Association for Medical Informatics, Biometry and Epidemiology; ICD, International Statistical Classification of Diseases and Related Health Problems; IPSCC, International Pediatric Sepsis Consensus Conference; PICU, Pediatric Intensive Care Unit; PDMS, Patient Data Management System; RCT, Randomized Controlled Trials; SIRS, Systemic Inflammatory Response Syndrome; STARD, Standards for Reporting Diagnostic Accuracy Studies; SPIRIT, Standard Protocol Items: Recommendations for Interventional Trials.

ARTICLE SUMMARY

Strengths and limitations of this study

- Related studies reached successful results in the context of clinical decision-support systems (CDSS) for SIRS detection, but due to the study design, the reported results often do not reflect the usefulness of such systems in routine clinical care.
- We present an adjusted and novel approach for the design and the statistical analysis of diagnostic studies for CDSSs from a more routine clinical care perspective, because our study comprises
- (1) the validation of the clinical decision-support system in comparison to the assessment of two experienced clinicians by blinded chart review (gold standard) and
- (2) the comparison of the system's diagnostic accuracy with the diagnostic accuracy of real-time assessments by clinicians working in routine clinical care and manually evaluating patients.
- Although our study does not comprise specific evaluations of CDSS user acceptance, it is suited to present the potentials of CDSS use in routine clinical care and, thus, to foster the willingness to trust the system in future.

INTRODUCTION

The first definition of systemic inflammatory response syndrome (SIRS) and sepsis was made by the members of the "American College of Chest Physicians/Society of Critical Care Medicine Consensus Conference Committee" in 1992. SIRS was described as a "systemic inflammatory response to a variety of severe clinical insults", sepsis on the other hand was "the systemic response to infection" [1]. The criteria have been adapted to pediatric patients by the International Pediatric Sepsis Consensus Conference (IPSCC) in 2005 [2]. According to this, SIRS was present if the patient presented two or more of the defined age-dependent criteria (at least an abnormal core temperature or leukocyte count). The other criteria include abnormal heart rate and respiratory rate. According to the IPSCC, sepsis is defined as SIRS in the presence of or as a result of suspected or proven infection. Although the definition of SIRS is no longer taken into account for the sepsis diagnosis in adult medicine, it is still relevant in pediatric medicine. Pediatric patients with SIRS and sepsis are known to have a higher risk of morbidity and mortality [3-6]. SIRS in pediatric patients also causes a significantly prolonged stay in intensive care after cardiothoracic surgery [7] and entails an increased probability of organdysfunctions [8].

Independent of the underlying disease and together with SIRS and sepsis, organ dysfunction and failure represent the critical determinants of patient outcome in both adult and pediatric medicine. Prevention and rapid effective treatment of multi-organ dysfunction and failure is crucial for survival. An optimization of the diagnostic and therapeutic workflow is likely to have an immense impact on clinical outcome of critically-ill patients. In pediatric septic shock patients, every hour without appropriate treatment was associated with an increased chance of death by 40% [9]. Conclusively, early recognition and treatment of pediatric SIRS and sepsis are vital [10].

To assure that patients are treated with the best available approaches, evidence-based medicine [11] together with personal expertise represent the current *gold standard* of medical patient management. However, clinicians are often confronted with a stressful environment, which fosters decision-making with a lower quality than aspired [12]. This is particularly true for pediatric intensive care units (PICU), in which clinicians work under challenging conditions characterized by a high degree of dynamics, uncertainty and risk, time pressure and a vast amount of data. Altogether, these factors carry risk for medical errors and adverse effects on patient safety [13-15].

To tackle these challenges, clinicians can be supported by clinical decision-support systems (CDSS). The growing digitalization in medicine involves an immense amount of highly heterogeneous datasets carrying the potential to be valuable for other purposes than initially expected (*secondary use of data*); the design of systems that are able to efficiently reuse and assimilate routine data, and thereby making data meaningful for clinical care, is fostered. CDSS are shining examples for systems processing clinical and non-clinical data and delivering an added value by detecting diseases, recommending therapies or uncovering yet unknown disease patterns [16]. Particularly in sophisticated intensive care settings, the use of highly developed patient data management systems (PDMS) allow the continuous recording of multiple clinical parameters and make high quality data available.

In our previous work, we designed a rule-based and interoperable CDSS for the detection of SIRS in pediatric intensive care[17]. The CDSS is able to retrieve and evaluate dynamic facts as routinely and automatically measured parameters from the bedside monitors to detect SIRS episodes. However, only when used in routine clinical care, the benefits can be translated into clinical care. Consequently, an extensive evaluation of the system's diagnostic accuracy is needed to assure that users will trust the system. This need is even aggravated in our context, because we strive for (1) using an automatic SIRS-labelling to train machine learning algorithms, and (2) reaching a self-learning system able to continuously process data and optimize its algorithms when used in routine care (*Learning Healthcare System*) [18]. In our previous work, we describe this approach for CDSS design, in which we denote the conduction of such a study as **second** important step towards the vision of **c**ross-institutional **a**nd **d**ata-driven **d**ecision support for **i**ntensive care **e**nvironments (CADDIE2) [18].

Related studies already reached satisfying results [19-21]. However, due to the study designs, the reported results might not reflect the usefulness of the CDSS in routine clinical care. Often, the coding of diagnoses as ICD (International Statistical Classification of Diseases and Related Health Problems) is used as gold standard against which the system's results are compared to. However, in ICD, episodes of SIRS and sepsis are not documented detailed enough, e.g. the time of occurrence is not described. Even though sometimes additional scores are used, not all relevant SIRS episodes can be reflected. Hence, systems evaluated with such an approach in fact have been successfully trained, but with respects to the ICD documentation and not routine decision-making. Additionally, the study population is often very preselected and requires a complete data set of all parameters required as input for the algorithm used. Such perfect data sets are often not available in clinical routine settings.

The exploration of factors influencing a successful CDSS implementation is a ubiquitous topic. In a recent literature review, Kilsdonk et al. [22] identified such factors for guideline-based CDSS implementation. One of the aspects reported mostly deals with the *information quality of the system* and covers the relevance of data and messages delivered by the system [22]. This finding relates to a well-known and obviously still unsolved issue called "alert fatigue" [23]. Other recent work published by Liberati et al. [24] describe the conduction of a qualitative study to identify different clusters of attitudes and barriers towards CDSS implementation. The authors describe that, together with a poor integration in the clinical workflow, the "fear of experiencing excessive number of alerts" [24] is one of the factors hindering the willingness to trust systems and to believe in their unforeseen opportunities (*mutual adjustment*). This step is declared as one of the most challenging obstacles in CDSS adoption. It is suggested to integrate and evaluate the CDSS in routine clinical care and with real users to overcome these limitations [24].

Against the background of our CADDIE objectives and the findings on successful CDSS implementations, we conclude that there is a need for a CADDIE2 trial focusing on validating the CDSS for SIRS and sepsis detection in routine clinical care.

Study objectives and diagnostic approaches

The **primary goal** is the evaluation of the diagnostic accuracy of the CDSS for detecting SIRS in PICU patients (*diagnostic approach I*), in comparison to the assessments of two clinicians by blinded chart review. In case of disagreement, a third clinician will be consulted. The expert assessments will be treated as gold standard and contain retrospective, extensive data analyses. These comprise evaluating all patients' measurements, not restricted to the SIRS parameters, including additional values for vital signs validated hourly by the attending nurse.

The **secondary goal** is to compare the diagnostic accuracy of the CDSS for detecting SIRS in PICU patients evaluated against this gold standard, to the diagnostic accuracy of routine assessments of different clinicians working in routine clinical care (*diagnostic approach II*) when compared to the same gold standard.

Trial design and study setting

The CADDIE2 study is designed as a single-arm, controlled, prospective diagnostic accuracy study. Single study center is the Department of Pediatric Cardiology and Intensive Care Medicine at the Hannover Medical School (monocentric). The estimated study duration is one year. Our study does not contain comparisons between different patient populations (single-arm) or interventions (no randomization). Each patient will be assessed by both diagnostic approaches.

METHODS AND ANALYSIS

Preceding studies

We can take advantage of the results of a randomized controlled trial (RCT) with 807 PICU patients from the same ward for the planned study. The expected SIRS prevalence on admission to PICU was reported as 5/100; 20-10% of the patients developed SIRS later on during their PICU stay [25]. Furthermore, we conducted a proof-of-concept study focusing on the technical practicability of the CDSS, yielding at promising results for both the technical infrastructure and the accuracy of the system (sensitivity of 1.00, specificity of 0.94) [17].

Recommendations and guidelines

We reviewed work on study planning, national recommendations and templates of ethics committees and associations, and followed the Good Clinical Practice (GCP) in non-drug trials. We designed our study in accordance with the "Standards for Reporting Diagnostic Accuracy Studies" (STARD) [26] and the "Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) [27] guidelines (see Figure 1, see additional file 1).

Patient population and eligibility criteria

All pediatric patients aged 0 to 18 years admitted to the study center - independent of the gender, the underlying disease or the time of admission - will be asked for their consent to participate. Patients will be recruited continuously and included, if a positive consent is available; and their length of stay exceeds twelve hours because any patient developing SIRS will not be discharged earlier.

The physicians are specialized pediatricians with experience in pediatric intensive care. There are always one experienced (working in this PICU for over a year) and one less experienced physician (working in this PICU for less than a year) in charge. The reviewers who will perform the manual chart review for creating gold standard decisions are specialized pediatricians and very experienced (working in this PICU for over three years), able to discriminate unsound and missing data.

Outcome measures

Sensitivity and specificity *on the level of patients* will be used as primary outcome measures. As second outcome measure, sensitivity and specificity also can be determined *on the level of intensive care days*.

Statistical analysis and sample size calculation

For the primary outcome measure, sensitivity and specificity will be determined together with *Wald confidence intervals*. The comparison will be carried out by comparing the lower bound of the confidence interval with the null hypothesis (which is, as described below in the sample size calculation paragraph, a sensitivity of 0.90, and a specificity of 0.80). If the lower bound of the 95% confidence intervals for sensitivity and specificity are both above the values of the pre-defined null hypotheses, we will reject the null hypotheses. For the secondary outcome measure, sensitivity and specificity will be determined together with confidence intervals based on *general estimating equations*. Additionally, for the secondary goal of comparing the diagnostic accuracy of the CDSS to the one of routine decisions (both when evaluated against the gold standard), sensitivity and specificity values will be compared by means of *McNemar tests* and confidence intervals constructed based on *general estimating equations*.

All analyses will be accompanied by secondary subgroup analyses, stratified e.g. by patients' age, type of shift and clinical picture associated with SIRS detection (including SIRS, sepsis, severe sepsis and septic shock). Factors that might modify the diagnostic accuracy of the CDSS will thus be evaluated in an exploratory way, allowing a better understanding of potential limitations of the system. SIRS prevalence and incidence will be monitored throughout the pilot phase and the main phase of the study, and will be compared to pre-study values in order to estimate the risk of a training effect on physicians' real-time diagnoses caused by knowledge about the aims of this study.

For analyzing the primary outcome measure, the assessment is carried out on the patient level. This is challenging since the assessment is not cross-sectional (as e.g. if the unit of assessment would be an hour respectively a shift) but needs to incorporate the complex longitudinal course of potential assessments within one patient. It is, however, the clinically most meaningful and the most conservative approach for estimating the diagnostic accuracy if conducted correctly. In our case, the entire period of stay is considered and information are aggregated on the patient level. Every person contributes (given that a correct diagnosis is restricted to the period of an hour respectively a shift) parts of its period of stay to the calculation of specificity independently of if the gold standard recorded a SIRS at some point since everybody will have periods without SIRS diagnosis (which need to be classified as well correctly by the CDSS). This leads to situations, which cannot be represented in only one cell of a contingency table. The classical four cases are amended by a new case, which occurs because the CDSS should assess the occurrence of a SIRS event with a correct timing (e.g. SIRS event is identified within the correct hour respectively shift). For example, this fifth case prevents that alert firings on day 30 of the intensive care stay will be evaluated as *true positives* if the gold standard reports a SIRS episode on day 2. Here, the CDSS did not identify the SIRS episode within the correct timing. Thus, this case is used for the determination of both *false positives* (day 30, contributing to specificity) and *false negatives* (day 2, contributing to sensitivity). Hence, the fifth case (false positive and false negative) can be defined as follows: *the gold standard reports at least one SIRS episode, and the CDSS detect SIRS episodes but (at least one) not within the same hour respectively shift*. All other cases are defined as usual (e.g. false positive: *the gold standard reports no SIRS episode but the CDSS detects one or multiple SIRS episodes*).

Based on the different cases, the sensitivity and the specificity will be determined independently. For sample size calculation, the results of the proof-of-concept study were used as a basis (sensitivity 1.00 and specificity 0.94, when calculating on the level of days). For this study with the modified statistical analysis approach, a sensitivity of 90% (alternative hypothesis: 0.98, null hypothesis: 0.90) and a specificity of 80% (alternative hypothesis: 0.90, null hypothesis: 0.80) were chosen as a clinically relevant diagnostic accuracy, with a given accuracy of estimate of 95% (type I error = 0.05) and a power of 90% (chi square test). Consequently, 97 patients suffering from at least one SIRS episode (for the estimation of sensitivity) as well as 137 patients with or without SIRS episodes are required (for the

estimation of specificity). Based on the expected incidence and prevalence, at least 300 patients need to be considered.

Timeline

Before study start, the clinicians were introduced in their tasks. No interventions, treatments or other care-related actions are prohibited and patients are treated with standard procedures (including data collection and measurements). Personal briefings on the routine documentation were carried out during this *pilot phase* (July 1, 2018, estimated duration: 1 month, see Figure 1). A designated physician will present the study to the patients, their parents or their legal guardians and ask for consent within the *recruiting phase* (August 1, 2018, estimated duration: 6 months). Simultaneously, clinicians will report their findings during their working shift per patient (routine assessments, *diagnostic approach II*). The clinicians do not perform extensive analyses of documentations or reported data (*assessment phase I*, August 1, 2018, estimated duration: 6 months). In the routine assessments it is documented whether the patient suffered from SIRS, sepsis or organ dysfunction (via digital documentation form, see Figure 2). The first two weeks of this phase will be treated as test phase.

Later on, two experienced clinicians will start with their weekly, extensive, blinded review and the definition of gold standard assessments per patient and per hour (*assessment phase II*, February 1, 2019, estimated duration: 3 months). As soon as 97 patients suffering from one or more SIRS episodes as well as 137 patients with one or without SIRS episodes have been identified, the recruitment will be terminated. Simultaneously, the data sets from all recruited patients will be integrated into a data repository to make them accessible for the CDSS. The CDSS assessments per patients and per hour will start (*diagnostic approach I*). In the final *analysis phase* (May 1, 2019, estimated duration: 2 months), the diagnostic accuracy of the CDSS will be evaluated by comparing the assessments to the gold standard decisions from the experts (primary goal of the study). Additionally, the diagnostic accuracy of the routine assessments will be determined by comparing them to the same gold standard decisions. The different accuracies can be compared (secondary goal of the study). Finally, the study results will be communicated via publication in a peer-reviewed journal.

Recruitment and consent

Eligible patients, their parents or their legal guardians will receive an information letter and a consent form (available in German, English, Turkish and Arabic) during a personal discussion with a physician. Additional information sheets for younger patients are available, one for children aged six to eleven and one for children aged 12 to 18. The families will receive privacy statement forms (data protection, accessibility and confidentiality; see additional file 2).

Patient involvement

There were no funds or time allocated for patient and public involvement so we were unable to involve patients. We intend to disseminate the results to the participants and will invite patients to help us developing an appropriate method of dissemination.

Data management and collection

The CDSS is an application with interfaces to a data repository, which is based on an semantic interoperability standard for clinical information representation (*openEHR* [28]). For more information, we refer to [17]. For the routine assessment (assessment phase I, diagnostic approach II), we created a documentation form, which is based on the same interoperability standard and the same interfaces to the data repository (see Figure 2). Thereby, all results (gold standard, diagnostic approach I "CDSS", diagnostic approach II "routine assessments") will be available in the same format. Patient data (identification, birthdate), intensive care parameters (heart rate, respiratory rate, body temperature, leukocytes, neutrophils, mechanical ventilation, cooling devices, pacemaker), alert parameters (SIRS, sepsis, organ dysfunction with duration, beginning and end of the episode and shift), and general documentations of the patient conditions, events or unintended effects will be documented and processed.

Data monitoring and auditing

Quality assurance measures are continuously carried out. Plausibility checks will be executed while integrating data into our repository (e.g. simple counts to can uncover whether data from the primary source is missing in the repository). Furthermore, the data set will be reviewed by physicians with respect to randomly selected observations to guarantee the

plausibility from a clinical perspective. By following the openEHR standard, we are able to automatically execute validation checks to uncover missing or wrong values when integrating the data sets or filling out the documentation form (e.g. definition of ranges for specific values, or double data entries). The study procedures will be monitored by the authors as well as by designated physicians and nurses. They will supervise the adherence to the study protocol, the procedures for routine documentation and data integration, data quality and privacy.

Data protection: data access and confidentiality

We designed a data protection concept in cooperation with the local data security officer. The concept defines pseudonymization and data access procedures, outlines the patient consent, and explains technical security mechanisms. All data sets collected or created as part of this study are treated as strictly confidential. The data sets will be stored pseudonymized and in secure conditions in the data repository located in the network of the Hannover Medical School. To prevent unauthorized disclosure of patient information, it is only accessible for the physicians and employees in charge of this study. Collected data from patients withdrawing their consent (drop outs), will be completely deleted from the data repository. All study files, the final study data sets as well as the study results will be archived for ten years in an approved long-term repository and in accordance to the relevant legal and statutory requirements. The patient will be informed about these procedures as well as their rights (including the possibility to withdraw the consent and to obtain information about collected data sets at any time), and will be asked to consent to these.

Ethics and dissemination

All aspects are designed according to the General Data Protection Regulation from the European Union (2016/679) and are accepted by the data security officer of the Hannover Medical School. A positive vote of the ethics committee was given (No. 7804_BO_S_2018, see additional file 3). Further details on data protection aspects can be requested from the authors. Results of the main study will be communicated via publication in a peer-reviewed journal. We intend to disseminate the results to the participants through an appropriate method of dissemination to be defined.

DISCUSSION

To be used in the long run, CDSSs have to deliver relevant information in a timely manner and at an adequate frequency. Current approaches for evaluating the usefulness of CDSS indeed present positive results. However, due to a restricted study design not designed towards daily work conditions, results may not represent the system feasibility in routine clinical care. With our work, we contribute a modified study design for evaluating the diagnostic accuracy of a CDSS with a strong focus on routine clinical care. We hypothesize that such an evaluation will demonstrate the potentials of CDSS use in routine clinical care. In case of a positive study outcome, we will be able to reason that our CDSS is not only feasible from a technical but also from a clinical perspective as it supports clinicians in critical diagnostic decision-making. For evaluation, a so-called gold standard representing the true state of the patient is required. However, in complex, knowledge-and experience-based contexts as diagnostic decision-making, reproducible, objective and quantitative "gold standards" are rare. We use an excessive evaluation of the patient data by two experience clinicians as benchmark. To reduce possible biases, the clinicians are blinded to each other and to the CDSS. In situations of disagreement, a third clinician will be consulted, decisions will be revealed and a consensus decision will be reported. We are aware that our approach is time-consuming requiring highly engaged clinicians. Because of the stressful PICU environment, assessments may be delayed, and thus, the study timeline may not be adhered. For an early recognition of issues and study monitoring at the ward, an assistant physician is in charge. Also the routine assessments of clinicians have to be managed as they are at the same risk to be biased. Clinical documentation might be handled more meticulous at the beginning and more careless in the end of the study. To prevent the latter, the new documentation form was designed in cooperation with the users and integrated in the PDMS used daily. Together with the designated study monitor, this integration raised the satisfactory and the utilization rate of the form as well as the adherence to the study protocol.

For sample size calculation, it might be possible that the incidence was overestimated, so that in our settings more than 300 patients are needed to reach 97 patients suffering from at least one SIRS episode. The recruitment will be continuously aligned towards the number of recruited SIRS patients to be able to stop the recruitment as soon as the required number has been reached. Our expected values for sensitivity and specificity are rather conservative because we decided to primarily use an equally conservative statistical analysis approach. However, the expected values are treated as acceptable in clinical routine as the diagnostic

accuracy of the system will be over a critical minimum (and with respect to the aspired second goal of our study, even better than in clinical routine decision-making). At the same time, alert fatigue will be prevented because specificity is equally high. We are confident that our thoughts meet the need for an optimum balance between sensitivity and specificity, e.g. as reported by Coleman et al. [23]. Nevertheless, we will enhance our results with a more liberal analysis on the level of days.

Our study has been successfully started with recruitment according to this design and promises valuable results. When reaching a good diagnostic accuracy compared to the gold standard as well as advantages over the diagnostic accuracy of routine assessments, we are optimistic that our users are willing to trust and use the system in future. Moreover, this will allow the conduction of future studies as for example the evaluation of patient outcomes, user acceptability, or real-time performance of the system.

STATEMENTS

Author Contributions

AW was responsible for design and implementation of the presented clinical decision-support system and the outline of the study protocol, and has drafted the manuscript. TJ provided clinical expertise for the use case and the design of the underlying knowledge model, leaded

the proof-of-concept study and co-drafted the manuscript. AK was primarily responsible for the design of the statistical analysis, the sample size calculation and the authoring of the corresponding sections. BS and SM helped in the conception of the general study approach but especially for definition of goals and outcome measures, timing and patient recruitment; SM is responsible for patient recruitment and monitors the study at the ward. PB and MM provided clinical expertise for study design, revised the manuscript critically, and gave subject-specific advices as well as the final approval of the manuscript version to be published. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Funding

No funding to declare.

Availability of data and material

The datasets generated and/or analyzed during the current study are not publicly available due to data privacy and security matters of patients but are available from the corresponding author on reasonable request.

Ethical approval and patient consent

This diagnostic study received ethics approval from the Hannover Medical School Ethics Committee (approval number 7804_BO_S_2018). The trial is registered with ClinicalTrials.gov (NCT03661450).

Consent to participate will be given by the patients, by their parents or by their legal guardians by signing a study consent form. All aspects are designed according to the General Data Protection Regulation (EU) 2016/679 and are accepted by the data security officer of the Hannover Medical School as well as by the Ethics Committee.

Protocol modifications will require a formal amendment to the protocol which will be reported to the Hannover Medical School Ethics Committee for approval.

Acknowledgements

We want to thank our colleagues from the ZIMt department for Educational and Scientific IT Systems of the Hannover Medical School for supporting in matters of data access and integration.

REFERENCES

1 American College of Chest Physicians/Society of Critical Care Medicine Consensus Conference: definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. *Crit Care Med* 1992;20(6):864–74. PubMed PMID: 1597042.

2 Goldstein B, Giroir B, Randolph A. International pediatric sepsis consensus conference: Definitions for sepsis and organ dysfunction in pediatrics*. *Pediatr Crit Care Med* 2005;6(1):2–8. doi: 10.1097/01.PCC.0000149131.72248.E6.

3 Wolfler A, Silvani P, Musicco M, et al. Incidence of and mortality due to sepsis, severe sepsis and septic shock in Italian Pediatric Intensive Care Units: A prospective national survey. *Intensive Care Med* 2008;34(9):1690–97. doi: 10.1007/s00134-008-1148-y.

4 Cvetkovic M, Lutman D, Ramnarayan P, et al. Timing of death in children referred for intensive care with severe sepsis: Implications for interventional studies. *Pediatric Critical Care Medicine* 2015;16(5):410–17. doi: 10.1097/PCC.0000000000000385.

5 Creedon JK, Vargas S, Asaro LA, et al. Timing of Antibiotic Administration in Pediatric Sepsis. *Pediatr Emerg Care* 2018. doi: 10.1097/PEC.0000000000001663.

6 Kortz TB, Sawe HR, Murray B, et al. Clinical Presentation and Outcomes among Children with Sepsis Presenting to a Public Tertiary Hospital in Tanzania. *Front Pediatr* 2017;5:278. doi: 10.3389/fped.2017.00278.

7 Boehne M, Sasse M, Karch A, et al. Systemic inflammatory response syndrome after pediatric congenital heart surgery: Incidence, risk factors, and clinical outcome. *J Card Surg* 2017;32(2):116–25. doi: 10.1111/jocs.12879.

8 Ganjoo S, Ahmad K, Qureshi UA, et al. Clinical Epidemiology of SIRS and Sepsis in Newly Admitted Children. *Indian J Pediatr* 2015;82(8):698–702. doi: 10.1007/s12098-014-1618-x.

9 Han YY, Carcillo JA, Dragotta MA, et al. Early reversal of pediatric-neonatal septic shock by community physicians is associated with improved outcome. *Pediatrics* 2003;112(4):793–99. doi: 10.1542/peds.112.4.793.

10 Paul R. Recognition, Diagnostics, and Management of Pediatric Severe Sepsis and Septic Shock in the Emergency Department. *Pediatr Clin North Am* 2018;65(6):1107–18. doi: 10.1016/j.pcl.2018.07.012.

11 Sackett DL. Evidence-based medicine. *Semin Perinatol* 1997;21(1):3–5. doi: 10.1016/S0146-0005(97)80013-4.

12 Zavala AM, Day GE, Plummer D, et al. Decision-making under pressure: Medical errors in uncertain and dynamic environments. *Aust Health Rev* 2017. doi: 10.1071/AH16088.

13 Lighthall GK, Vazquez-Guillamet C. Understanding Decision Making in Critical Care. *Clin Med Res* 2015;13(3-4):156–68. doi: 10.3121/cmr.2015.1289.

14 Castaneda C, Nalley K, Mannion C, et al. Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine. *J Clin Bioinforma* 2015;5:4. doi: 10.1186/s13336-015-0019-3.

15 Lincoln P, Manning M-J, Hamilton S, et al. A pediatric critical care practice group: use of expertise and evidence-based practice in identifying and establishing "best" practice. *Crit Care Nurse* 2013;33(2):85–87. doi: 10.4037/ccn2013740.

16 Shortliffe EH. Biomedical Informatics: Defining the Science and Its Role in Health Professional Education. In: Hutchison D, Kanade T, Kittler J, et al., eds. Information Quality in e-Health. Berlin, Heidelberg: Springer Berlin Heidelberg 2011:711–14.

17 Wulff A, Haarbrandt B, Tute E, et al. An interoperable clinical decision-support system for early detection of SIRS in pediatric intensive care using openEHR. *Artif Intell Med* 2018;89:10–23. doi: 10.1016/j.artmed.2018.04.012.

18 Wulff A, Marschollek M. Learning Healthcare Systems in Pediatrics: Cross-Institutional and Data-Driven Decision-Support for Intensive Care Environments (CADDIE). *Stud Health Technol Inform* 2018;251:109–12. doi:10.3233/978-1-61499-880-8-109

19 Faisal M, Scally A, Richardson D, et al. Development and External Validation of an Automated Computer-Aided Risk Score for Predicting Sepsis in Emergency Medical Admissions Using the Patient's First Electronically Recorded Vital Signs and Blood Test Results. *Crit Care Med* 2018;46(4):612–18. doi: 10.1097/CCM.0000000000002967.

20 Austrian JS, Jamin CT, Doty GR, et al. Impact of an emergency department electronic sepsis surveillance system on patient mortality and length of stay. *J Am Med Inform Assoc* 2018;25(5):523–29. doi: 10.1093/jamia/ocx072.

21 Mao Q, Jay M, Hoffman JL, et al. Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department, general ward and ICU. *BMJ Open* 2018;8(1):e017833. doi: 10.1136/bmjopen-2017-017833.

22 Kilsdonk E, Peute LW, Jaspers MW. Factors influencing implementation success of guideline-based clinical decision support systems: A systematic review and gaps analysis. *Int J Med Inform* 2017;98:56–64. doi: 10.1016/j.ijmedinf.2016.12.001.

23 Coleman JJ, van der Sijs H, Haefeli WE, et al. On the alert: Future priorities for alerts in clinical decision support for computerized physician order entry identified from a European workshop. *BMC Med Inform Decis Mak* 2013;13:111. doi: 10.1186/1472-6947-13-111.

24 Liberati EG, Ruggiero F, Galuppo L, et al. What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. *Implement Sci* 2017;12(1):113. doi: 10.1186/s13012-017-0644-2.

25 Jack T, Boehne M, Brent BE, et al. In-line filtration reduces severe complications and length of stay on pediatric intensive care unit: A prospective, randomized, controlled trial. *Intensive Care Med* 2012;38(6):1008–16. doi: 10.1007/s00134-012-2539-7.

26 Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: Explanation and elaboration. *BMJ Open* 2016;6(11):e012799. doi: 10.1136/bmjopen-2016-012799.

27 Chan A-W, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: Guidance for protocols of clinical trials. *BMJ* 2013;346:e7586. doi: 10.1136/bmj.e7586.

28 Thomas Beale. Archetypes: Constraint-based Domain Models for Future-proof Information Systems. In: Eleventh OOPSLA workshop on behavioral semantics: serving the customer. Seattle, Washington, Boston: Northeastern University 2002:16–32.

Figure 1: Time schedule and study episodes for CADDIE2 trial

Figure 2: Digital documentation form (based on openEHR data repository)

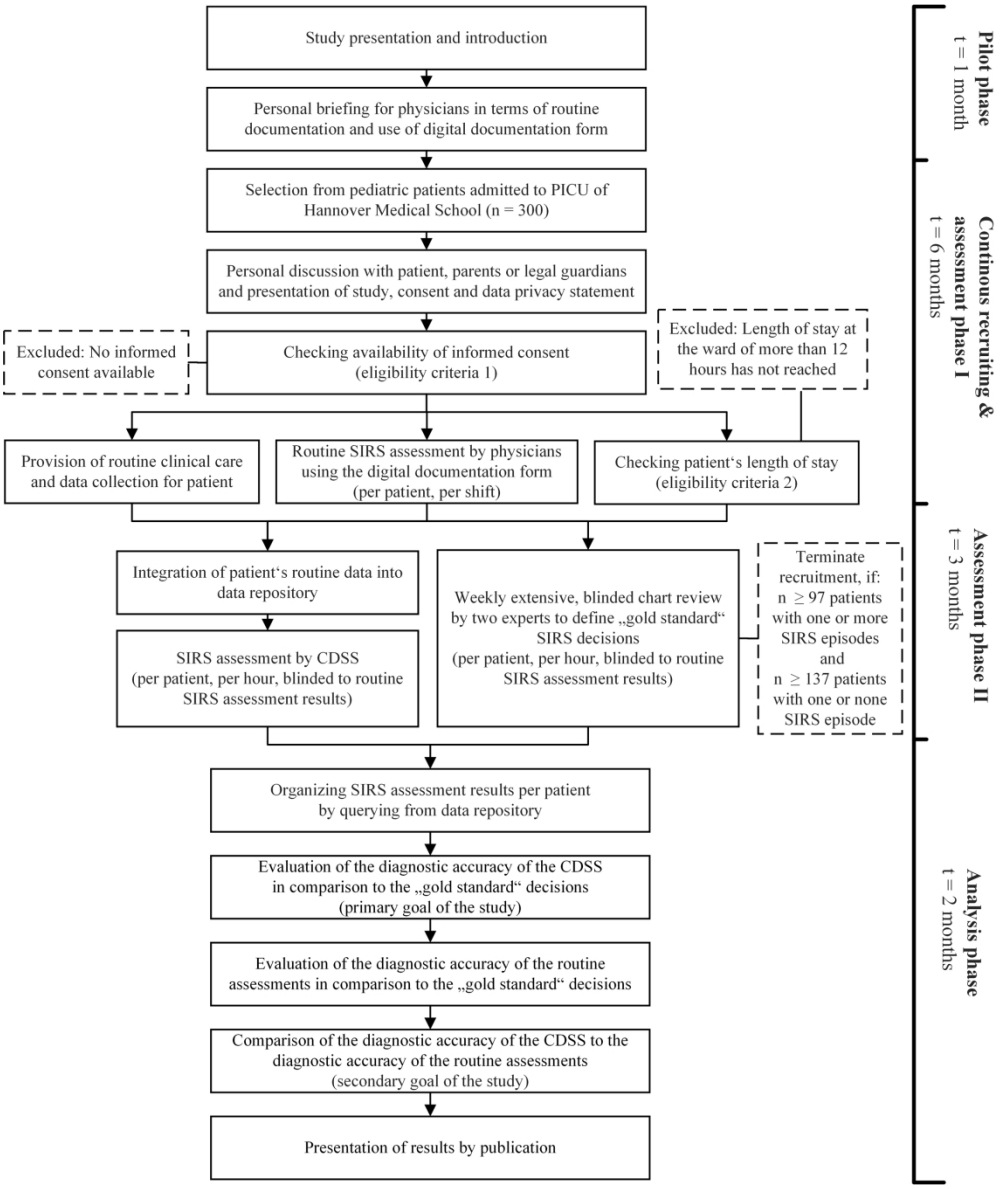

Figure 1: Time schedule and study episodes for CADDIE2 trial

208x245mm (300 x 300 DPI)

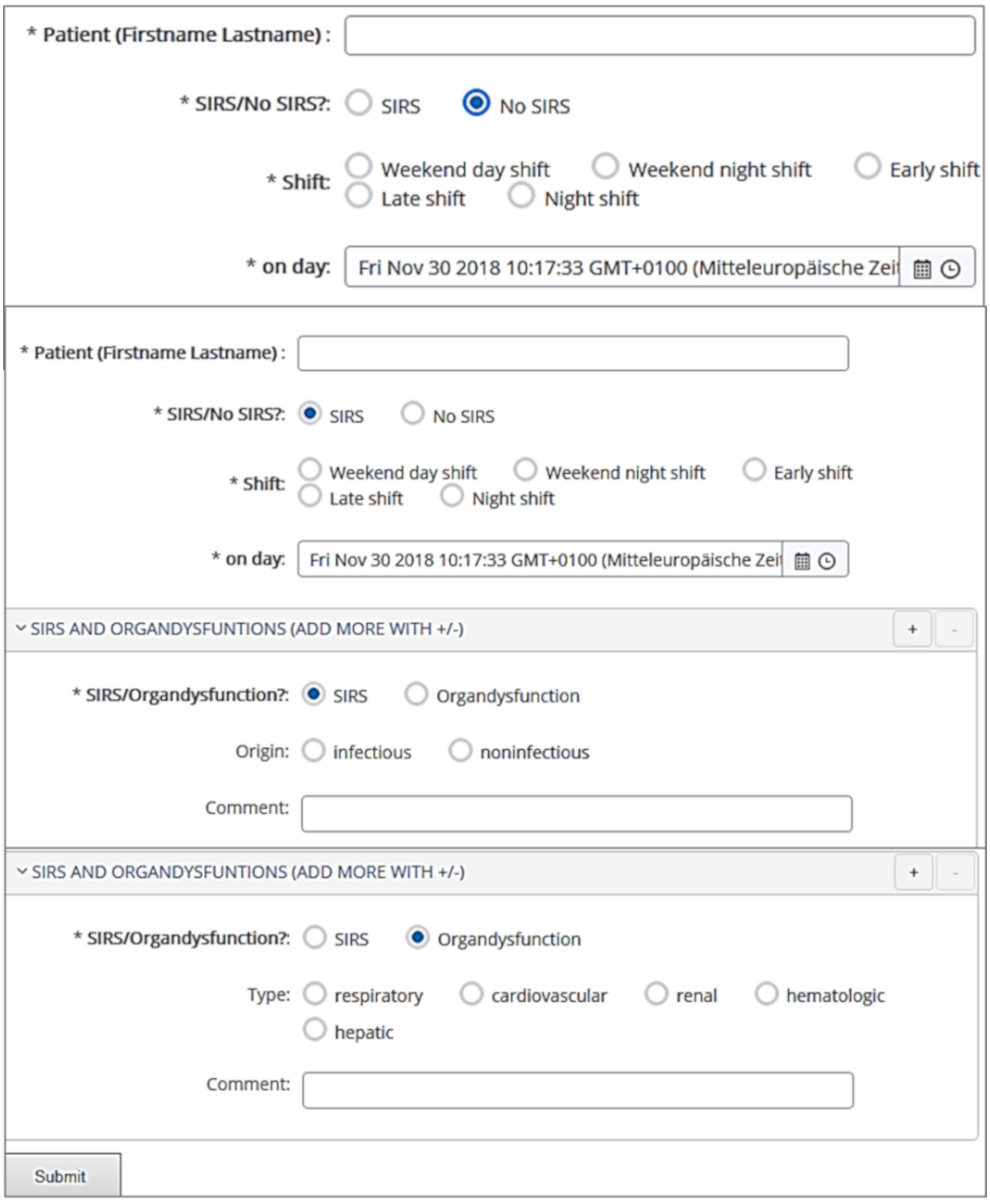

Figure 2: Digital documentation form (based on openEHR data repository)

90x109mm (300 x 300 DPI)

SPIRIT

STANDARD PROTOCOL ITEMS: RECOMMENDATIONS FOR INTERVENTIONAL TRIALS

SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*

Section/item	Item No	Description	Addressed on page number
Administrative information			
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	1
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry	3
	2b	All items from the World Health Organization Trial Registration Data Set	NA
Protocol version	3	Date and version identifier	3
Funding	4	Sources and types of financial, material, and other support	18
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	1,18
	5b	Name and contact information for the trial sponsor	NA
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	18
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	9,15,16

Introduction			
Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	_____5,6,7__
	6b	Explanation for choice of comparators	_____5,6,7,8__
Objectives	7	Specific objectives or hypotheses	_____7,8_____
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)	_____8_____
Methods: Participants, interventions, and outcomes			
Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	_____8_____
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	_____9_____
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	NA (no interventions, diagnostic study), alternative: _____8_____
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)	_____NA (each patient will be assessed by both diagnostic approaches), alternative: _____8_____

	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)	_____NA (no intervention, routine care), alternative: _____16____
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	_____10____
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	_____9,10___
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	_____10,11,12_
Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	_____9,10____
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size	_____12,13__

Methods: Assignment of interventions (for controlled trials)

Allocation:

Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	_____NA____
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	_____NA____
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	_____NA____

Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	7,11,16
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	16
Methods: Data collection, management, and analysis			
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	13,14
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	NA (no interventions)
Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	13,14, 15
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	9,10
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	NA
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	NA
Methods: Monitoring			
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	15
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	NA

Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	15
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	15,16
Ethics and dissemination			
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval	16, 18, 19, Additional file 3
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)	19
Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	12,13
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	NA
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	13,15
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site	18
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	18
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	NA
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	11,16

	31b	Authorship eligibility guidelines and any intended use of professional writers	_____18____
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	_____NA___
Appendices			
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates	_____Additional file 2__
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	_____NA____

*It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items. Amendments to the protocol should be tracked and dated. The SPIRIT checklist is copyrighted by the SPIRIT Group under the Creative Commons "Attribution-NonCommercial-NoDerivs 3.0 Unported" license.

Patient Information (Parents)

Dear parents,

we would like to include your child in a clinical trial, you will find the details on the following pages. Please read them carefully. If you have any questions please do not hesitate to ask the doctor in charge.

Information about the clinical trial which is named:

Evaluation of the diagnostic accuracy of a Clinical Decision-Support System (CDSS) to support recognition of SIRS and Sepsis in paediatric intensive care patients in comparison to medical specialists.

Your child was or will be admitted on the paediatric intensive care unit (PICU) at Medizinische Hochschule Hannover. Due to neccessary monitoring and therapy we will gather laboratory results and vital signs continueally.

During the clinical trial these laboratory and monitoring details will be analyzed with regard to a pronounced, generalized inflammatory reaction (Systemic inflammatory response syndrome (SIRS)) using the support of a computer-based system. The objective is to support the doctors in early recognition of SIRS.

The recognition of SIRS on a PICU is a complex task that is quite suitable for a Clinical Decision-Support System. SIRS is caused by different reasons, for example surgery, injurys or infection. It is a typical occurence at PICU-patients and puts the patient under additional strain independent from the original reason for admittance. For diagnosing SIRS in paediatric patients correctly we have to find abnormalities in different vital signs and laboratory results which have age related reference values, making it an advanced task.

The development and clinical trial of a CDSS for SIRS in paediatric patients is the result of a cooperation between the Department of Paediatric Cardiology and Intensive Care Medicine of Medizinische Hochschule Hannover and the Peter L. Reichertz Institut für

Medizinsche Informatik der Technischen Universität Braunschweig und der Medizinischen Hochschule Hannover (PLRI).

Are there any risks by taking part in the clinical trial?

The CDSS will only analyze data that is taken and saved during your child's admittance on the PICU anyway. There won't be any additional blood tests, monitoring or examinations. These standardized raised data will be saved in a pseudonymized way that ensures the patient's privacy.

What happens if I don't agree to take part in the clinical trial?

The whole team will respect your decision if you do not want your child to take part in the clinical trial. This will not be evaluated or have any negative effect on your child's treatment. You can also revoke your agreement anytime.

Privacy Policy

The data raised and saved for the trial are pseudonymized and it is impossible to reconnect them to an individual. The data will be saved for 10 years and deleted afterwards. The issues of the Bundesdatenschutzgesetz are considered in every respect.

I received the information and consent to the clinical trial mentioned above.

☐Yes ☐No

_____

(Date and signature of the parent/legal guardian)

_____

(Date and signature of the medical doctor)

Patient Information (Patient)

Dear patient,

we would like to include you in a clinical trial, you will find the details on the following pages. Please read them carefully. If you have any questions please do not hesitate to ask the doctor in charge.

Information about the clinical trial which is named:

Evaluation of the diagnostic accuracy of a Clinical Decision-Support System (CDSS) to support recognition of SIRS and Sepsis in paediatric intensive care patients in comparison to medical specialists.

You were or will be admitted on the paediatric intensive care unit (PICU) at Medizinische Hochschule Hannover. To treat you best we will monitor your vital signs as heartbeat, blood pressure and respiratory rate and take blood tests to monitor your white blood count and other inflammatory parameters. The results will be saved.

We know about the occurence of extreme inflammatory reactions in paediatric patients. These can be cause by several reasons as surgery, injury or infection. We can measure these by changes in your vital signs, temperature or white blood count.
With a clinical trial something new is tested, for example some treatment. For this clinical trial your monitoring or vital sign data will be analyzed by a computer programme to support your doctors to detect changes that might be caused by an inflammatory reaction as soon as possible. The sooner we are aware we can adjust your treatment.

Are there any risks by taking part in the clinical trial?

The CDSS will only analyze data that is taken and saved during your admittance on the PICU anyway. There won't be any additional blood tests, monitoring or examinations. These standardized raised data will be saved in a pseudonymized way that ensures the patient's privacy.

What happens if I don't agree to take part in the clinical trial?

The whole team will respect your decision if you do not want to take part in the clinical trial. This will not be evaluated or have any negative effect on your treatment. You can also revoke your agreement anytime.

Privacy Policy

The data raised and saved for the trial are pseudonymized and it is impossible to reconnect them to an individual. The data will be saved for 10 years and deleted afterwards. The issues of the Bundesdatenschutzgesetz are considered in every respect.

I received the information and consent to the clinical trial mentioned above.

☐Yes ☐No

_____

(Date and signature of the patient)

_____

(Date and signature of the medical doctor)

Medizinische Hochschule Hannover

MHH Ethikkommission OE 9515
30623 Hannover

Herrn
Dr. Thomas Jack
Klinik für Pädiatrische Kardiologie und Pädiatrische Intensivmedizin
OE 6730 – im Hause

Ethikkommission
Vorsitzender:
Prof. Dr. Stefan Engeli

Sekretariat:
Marion Lange
Telefon: 0511 532-3443
Liane Höft
Telefon: 0511 532-9812

Fax: 0511 532-16 3443
ethikkommission@mh-hannover.de

Carl-Neuberg-Straße 1
30625 Hannover
Telefon: 0511 532-0
www.mh-hannover.de

20.06.2018/MLa

Nr. 7804_BO_S_2018
Evaluation der diagnostischen Genauigkeit eines Clinical Decision-Support Systems (CDSS) zur Unterstützung der Erkennung von SIRS und Sepsis auf der pädiatrischen Intensivstation im Vergleich zu medizinischem Fachpersonal

Sehr geehrter Herr Kollege Jack,

die Mitglieder der Ethikkommission haben auf ihrer Sitzung am 21.03.2018 über o.g. Antrag beraten. Nach Eingang der gemäß unserem Schreiben vom 26.03.2018 überarbeiteten Unterlagen bestehen keine Bedenken gegenüber der Durchführung der Studie.

Die Ethikkommission weist darauf hin, dass die ärztliche und juristische Verantwortung bei den jeweiligen Prüfärzten verbleibt.

Datenschutzrechtliche Aspekte von Forschungsvorhaben werden durch die Ethikkommission grundsätzlich nur kursorisch geprüft. Dieses Votum / diese Bewertung ersetzt mithin nicht die Konsultation des zuständigen Datenschutzbeauftragten.

Mit besten Grüßen

Prof. Dr. Stefan Engeli
Vorsitzender der Ethikkommission

Folgende Mitglieder haben an der Beratung des o.g. Antrages mitgewirkt:

Prof. Dr. St. Engeli (Vorsitzender), Klinische Pharmakologie der MHH
PD Dr. U.-V. Albrecht, Stellv. Leiter des Peter L. Reichertz Institut für Medizinische Informatik
Prof. Dr. A. M. Das, Leiter der Pädiatrischen Stoffwechselmedizin, Klinik für Päd. Nieren-, Leber- und Stoffwechselerkrankungen, Zentrum Kinderheilkunde und Jugendmedizin der MHH
Dr. J. Graubner, Arzt für Allgemeinmedizin
Prof. Dr. Thomas Illig, Leiter Biobank (HUB)
Prof. Dr. A. Koch, Leiter der Abt. Biometrie der MHH
Frau Prof. Dr. B. Lohff, Em. Leiterin der Abt. Geschichte, Ethik und Philosophie der Medizin der MHH
Dr. Oliver Pramann, Fachanwalt für Medizinrecht
Jörg Viering, Leiter der Forschungswerkstätten, MHH
Prof. Dr. Peter Vogt, Leiter der Klinik für Plastische und Wiederherstellungschirurgie, MHH

BMJ Open

CADDIE2 - Evaluation of a clinical decision-support system for early detection of systemic inflammatory response syndrome in pediatric intensive care: study protocol for a diagnostic study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-028953.R2
Article Type:	Protocol
Date Submitted by the Author:	22-May-2019
Complete List of Authors:	Wulff, Antje; Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School Montag, Sara; Department of Pediatric Cardiology and Intensive Care Medicine, Hannover Medical School Steiner, Bianca; Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School Marschollek, Michael; Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School Beerbaum, Philipp; Department of Pediatric Cardiology and Intensive Care Medicine, Hannover Medical School Karch, André; Institute of Epidemiology and Social Medicine, University of Muenster Jack, Thomas; Department of Pediatric Cardiology and Intensive Care Medicine, Hannover Medical School
Primary Subject Heading:	Paediatrics
Secondary Subject Heading:	Health informatics, Intensive care, Infectious diseases, Diagnostics
Keywords:	Clinical Decision Support Systems, Clinical Trial, Systemic Inflammatory Response Syndrome, Paediatric intensive & critical care < INTENSIVE & CRITICAL CARE

CADDIE2 - Evaluation of a clinical decision-support system for early detection of systemic inflammatory response syndrome in pediatric intensive care: study protocol for a diagnostic study

Antje Wulff[a,1], Sara Montag[b], Bianca Steiner[a], Michael Marschollek[a], Philipp Beerbaum[b], André Karch[c], Thomas Jack[b]

[a]*Peter L. Reichertz Institute for Medical Informatics of TU Braunschweig and Hannover Medical School, Braunschweig and Hannover, Germany*

[b]*Department of Pediatric Cardiology and Intensive Care Medicine, Hannover Medical School, Hannover, Germany*

[c]*Institute of Epidemiology and Social Medicine, University of Muenster, Muenster, Germany*

[1]Corresponding author at:
Peter L. Reichertz Institute for Medical Informatics of
TU Braunschweig and Hannover Medical School,
Carl-Neuberg-Str. 1, 30625 Hannover, Germany.
E-mail address: antje.wulff@plri.de (A.Wulff).
URL: http://www.plri.de (A. Wulff).

Email addresses for all authors:
antje.wulff@plri.de
montag.sara@mh-hannover.de
bianca.steiner@plri.de
michael.marschollek@plri.de
beerbaum.philipp@mh-hannover.de
andre.karch@ukmuenster.de
jack.thomas@mh-hannover.de

Word Count (excluding title page, abstract, references, figures and tables and captions):

4047

ABSTRACT

Introduction: Systemic inflammatory response syndrome (SIRS) is one of the most critical indicators determining the clinical outcome of pediatric intensive care patients. Clinical decision-support systems (CDSS) can be designed to support clinicians in detection and treatment. However, the use of such systems is highly discussed as they are often associated with accuracy problems and "alert fatigue". We designed a CDSS for detection of pediatric SIRS and hypothesize that a high diagnostic accuracy together with an adequate alerting will accelerate the use. Our study will (1) determine the diagnostic accuracy of the CDSS compared to gold standard decisions created by two, blinded experienced pediatricians, and (2) compare the system's diagnostic accuracy with that of routine clinical care decisions compared to the same gold standard.

Methods and analysis: CADDIE2 is a prospective diagnostic accuracy study taking place at the Department of Pediatric Cardiology and Intensive Care Medicine at the Hannover Medical School; it represents the second step towards our vision of cross-institutional and data-driven decision support for intensive care environments (CADDIE). The study comprises (1) recruitment of up to 300 patients (started *August 1, 2018*), (2) creation of gold standard decisions (start date *May 1, 2019*), (3) routine SIRS assessments by physicians (started with recruitment), (4) SIRS assessments by a CDSS (start date *May 1, 2019*), and (5) statistical analysis with a modified approach for determining sensitivity and specificity and comparing the accuracy results of the different diagnostic approaches (planned start date *July 1, 2019*).

Ethics and dissemination: Ethics approval was obtained at the study center (Ethics Committee of Hannover Medical School). Results of the main study will be communicated via publication in a peer-reviewed journal.

Trial registration: ClinicalTrials.gov NCT03661450. Registered 07 September 2018. Recruitment started August 1st, 2018, and it is expected to continue until May 2019.

Protocol version: V1.0, 2018-Mar-06 Original

Keywords: Clinical Decision Support Systems, Clinical Trial, Paediatric intensive & critical care, Systemic Inflammatory Response Syndrome

Abbreviations: CADDIE, Cross-institutional and Data-driven Decision Support for Intensive Care Environments; CDSS, Clinical Decision-Support System; GCP, Good Clinical Practice; GMDS, German Association for Medical Informatics, Biometry and Epidemiology; ICD, International Statistical Classification of Diseases and Related Health Problems; IPSCC, International Pediatric Sepsis Consensus Conference; PICU, Pediatric Intensive Care Unit; PDMS, Patient Data Management System; RCT, Randomized Controlled Trials; SIRS, Systemic Inflammatory Response Syndrome; STARD, Standards for Reporting Diagnostic Accuracy Studies; SPIRIT, Standard Protocol Items: Recommendations for Interventional Trials.

ARTICLE SUMMARY

Strengths and limitations of this study

- Related studies reached successful results in the context of clinical decision-support systems (CDSS) for SIRS detection, but due to the study design, the reported results often do not reflect the usefulness of such systems in routine clinical care.
- We present an adjusted and novel approach for the design and the statistical analysis of diagnostic studies for CDSSs from a more routine clinical care perspective, because our study comprises
- (1) the validation of the clinical decision-support system in comparison to the assessment of two experienced clinicians by blinded chart review (gold standard) and
- (2) the comparison of the system's diagnostic accuracy with the diagnostic accuracy of real-time assessments by clinicians working in routine clinical care and manually evaluating patients.
- Although our study does not comprise specific evaluations of CDSS user acceptance, it is suited to present the potentials of CDSS use in routine clinical care and, thus, to foster the willingness to trust the system in future.

INTRODUCTION

The first definition of systemic inflammatory response syndrome (SIRS) and sepsis was made by the members of the "American College of Chest Physicians/Society of Critical Care Medicine Consensus Conference Committee" in 1992. SIRS was described as a "systemic inflammatory response to a variety of severe clinical insults", sepsis on the other hand was "the systemic response to infection" [1]. The criteria have been adapted to pediatric patients by the International Pediatric Sepsis Consensus Conference (IPSCC) in 2005 [2]. According to this, SIRS was present if the patient presented two or more of the defined age-dependent criteria (at least an abnormal core temperature or leukocyte count). The other criteria include abnormal heart rate and respiratory rate. According to the IPSCC, sepsis is defined as SIRS in the presence of or as a result of suspected or proven infection. Although the definition of SIRS is no longer taken into account for the sepsis diagnosis in adult medicine, it is still relevant in pediatric medicine. Pediatric patients with SIRS and sepsis are known to have a higher risk of morbidity and mortality [3-6]. SIRS in pediatric patients also causes a significantly prolonged stay in intensive care after cardiothoracic surgery [7] and entails an increased probability of organdysfunctions [8].

Independent of the underlying disease and together with SIRS and sepsis, organ dysfunction and failure represent the critical determinants of patient outcome in both adult and pediatric medicine. Prevention and rapid effective treatment of multi-organ dysfunction and failure is crucial for survival. An optimization of the diagnostic and therapeutic workflow is likely to have an immense impact on clinical outcome of critically-ill patients. In pediatric septic shock patients, every hour without appropriate treatment was associated with an increased chance of death by 40% [9]. Conclusively, early recognition and treatment of pediatric SIRS and sepsis are vital [10].

To assure that patients are treated with the best available approaches, evidence-based medicine [11] together with personal expertise represent the current *gold standard* of medical patient management. However, clinicians are often confronted with a stressful environment, which fosters decision-making with a lower quality than aspired [12]. This is particularly true for pediatric intensive care units (PICU), in which clinicians work under challenging conditions characterized by a high degree of dynamics, uncertainty and risk, time pressure and a vast amount of data. Altogether, these factors carry risk for medical errors and adverse effects on patient safety [13-15].

To tackle these challenges, clinicians can be supported by clinical decision-support systems (CDSS). The growing digitalization in medicine involves an immense amount of highly heterogeneous datasets carrying the potential to be valuable for other purposes than initially expected (*secondary use of data*); the design of systems that are able to efficiently reuse and assimilate routine data, and thereby making data meaningful for clinical care, is fostered. CDSS are shining examples for systems processing clinical and non-clinical data and delivering an added value by detecting diseases, recommending therapies or uncovering yet unknown disease patterns [16]. Particularly in sophisticated intensive care settings, the use of highly developed patient data management systems (PDMS) allow the continuous recording of multiple clinical parameters and make high quality data available.

In our previous work, we designed a rule-based and interoperable CDSS for the detection of SIRS in pediatric intensive care[17]. The CDSS is able to retrieve and evaluate dynamic facts as routinely and automatically measured parameters from the bedside monitors to detect SIRS episodes. However, only when used in routine clinical care, the benefits can be translated into clinical care. Consequently, an extensive evaluation of the system's diagnostic accuracy is needed to assure that users will trust the system. This need is even aggravated in our context, because we strive for (1) using an automatic SIRS-labelling to train machine learning algorithms, and (2) reaching a self-learning system able to continuously process data and optimize its algorithms when used in routine care (*Learning Healthcare System*) [18]. In our previous work, we describe this approach for CDSS design, in which we denote the conduction of such a study as **second** important step towards the vision of **c**ross-institutional **a**nd **d**ata-driven **d**ecision support for **i**ntensive care **e**nvironments (CADDIE2) [18].

Related studies already reached satisfying results [19-21]. However, due to the study designs, the reported results might not reflect the usefulness of the CDSS in routine clinical care. Often, the coding of diagnoses as ICD (International Statistical Classification of Diseases and Related Health Problems) is used as gold standard against which the system's results are compared to. However, in ICD, episodes of SIRS and sepsis are not documented detailed enough, e.g. the time of occurrence is not described. Even though sometimes additional scores are used, not all relevant SIRS episodes can be reflected. Hence, systems evaluated with such an approach in fact have been successfully trained, but with respects to the ICD documentation and not routine decision-making. Additionally, the study population is often very preselected and requires a complete data set of all parameters required as input for the algorithm used. Such perfect data sets are often not available in clinical routine settings.

The exploration of factors influencing a successful CDSS implementation is a ubiquitous topic. In a recent literature review, Kilsdonk et al. [22] identified such factors for guideline-based CDSS implementation. One of the aspects reported mostly deals with the *information quality of the system* and covers the relevance of data and messages delivered by the system [22]. This finding relates to a well-known and obviously still unsolved issue called "alert fatigue" [23]. Other recent work published by Liberati et al. [24] describe the conduction of a qualitative study to identify different clusters of attitudes and barriers towards CDSS implementation. The authors describe that, together with a poor integration in the clinical workflow, the "fear of experiencing excessive number of alerts" [24] is one of the factors hindering the willingness to trust systems and to believe in their unforeseen opportunities (*mutual adjustment*). This step is declared as one of the most challenging obstacles in CDSS adoption. It is suggested to integrate and evaluate the CDSS in routine clinical care and with real users to overcome these limitations [24].

Against the background of our CADDIE objectives and the findings on successful CDSS implementations, we conclude that there is a need for a CADDIE2 trial focusing on validating the CDSS for SIRS and sepsis detection in routine clinical care.

Study objectives and diagnostic approaches

The **primary goal** is the evaluation of the diagnostic accuracy of the CDSS for detecting SIRS in PICU patients (*diagnostic approach I*), in comparison to the assessments of two clinicians by blinded chart review. In case of disagreement, a third clinician will be consulted. The expert assessments will be treated as gold standard and contain retrospective, extensive data analyses. These comprise evaluating all patients' measurements, not restricted to the SIRS parameters, including additional values for vital signs validated hourly by the attending nurse.

The **secondary goal** is to compare the diagnostic accuracy of the CDSS for detecting SIRS in PICU patients evaluated against this gold standard, to the diagnostic accuracy of routine assessments of different clinicians working in routine clinical care (*diagnostic approach II*) when compared to the same gold standard.

Trial design and study setting

The CADDIE2 study is designed as a single-arm, controlled, prospective diagnostic accuracy study. Single study center is the Department of Pediatric Cardiology and Intensive Care Medicine at the Hannover Medical School (monocentric). The estimated study duration is one year. Our study does not contain comparisons between different patient populations (single-arm) or interventions (no randomization). Each patient will be assessed by both diagnostic approaches.

METHODS AND ANALYSIS

Preceding studies

We can take advantage of the results of a randomized controlled trial (RCT) with 807 PICU patients from the same ward for the planned study. The expected SIRS prevalence on admission to PICU was reported as 5/100; 20-10% of the patients developed SIRS later on during their PICU stay [25]. Furthermore, we conducted a proof-of-concept study focusing on the technical practicability of the CDSS, yielding at promising results for both the technical infrastructure and the accuracy of the system (sensitivity of 1.00, specificity of 0.94) [17].

Recommendations and guidelines

We reviewed work on study planning, national recommendations and templates of ethics committees and associations, and followed the Good Clinical Practice (GCP) in non-drug trials. We designed our study in accordance with the "Standards for Reporting Diagnostic Accuracy Studies" (STARD) [26] and the "Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) [27] guidelines (see Figure 1, see additional file 1).

Patient population and eligibility criteria

All pediatric patients aged 0 to 18 years admitted to the study center - independent of the gender, the underlying disease or the time of admission - will be asked for their consent to participate. Patients will be recruited continuously and included, if a positive consent is available; and their length of stay exceeds twelve hours because any patient developing SIRS will not be discharged earlier.

The physicians are specialized pediatricians with experience in pediatric intensive care. There are always one experienced (working in this PICU for over a year) and one less experienced physician (working in this PICU for less than a year) in charge. The reviewers who will perform the manual chart review for creating gold standard decisions are specialized pediatricians and very experienced (working in this PICU for over three years), able to discriminate unsound and missing data.

Outcome measures

Sensitivity and specificity *on the level of patients* will be used as primary outcome measures. As second outcome measure, sensitivity and specificity also can be determined *on the level of intensive care days*.

Statistical analysis and sample size calculation

For the primary outcome measure, sensitivity and specificity will be determined together with *Wald confidence intervals*. The comparison will be carried out by comparing the lower bound of the confidence interval with the null hypothesis (which is, as described below in the sample size calculation paragraph, a sensitivity of 0.90, and a specificity of 0.80). If the lower bound of the 95% confidence intervals for sensitivity and specificity are both above the values of the pre-defined null hypotheses, we will reject the null hypotheses. For the secondary outcome measure, sensitivity and specificity will be determined together with confidence intervals based on *general estimating equations*. Additionally, for the secondary goal of comparing the diagnostic accuracy of the CDSS to the one of routine decisions (both when evaluated against the gold standard), sensitivity and specificity values will be compared by means of *McNemar tests* and confidence intervals constructed based on *general estimating equations*.

All analyses will be accompanied by secondary subgroup analyses, stratified e.g. by patients' age, type of shift and clinical picture associated with SIRS detection (including SIRS, sepsis, severe sepsis and septic shock). Factors that might modify the diagnostic accuracy of the CDSS will thus be evaluated in an exploratory way, allowing a better understanding of potential limitations of the system. SIRS prevalence and incidence will be monitored throughout the pilot phase and the main phase of the study, and will be compared to pre-study values in order to estimate the risk of a training effect on physicians' real-time diagnoses caused by knowledge about the aims of this study.

For analyzing the primary outcome measure, the assessment is carried out on the patient level. This is challenging since the assessment is not cross-sectional (as e.g. if the unit of assessment would be an hour respectively a shift) but needs to incorporate the complex longitudinal course of potential assessments within one patient. It is, however, the clinically most meaningful and the most conservative approach for estimating the diagnostic accuracy if conducted correctly. In our case, the entire period of stay is considered and information are aggregated on the patient level. Every person contributes (given that a correct diagnosis is restricted to the period of an hour respectively a shift) parts of its period of stay to the calculation of specificity independently of if the gold standard recorded a SIRS at some point since everybody will have periods without SIRS diagnosis (which need to be classified as well correctly by the CDSS). This leads to situations, which cannot be represented in only one cell of a contingency table. The classical four cases are amended by a new case, which occurs because the CDSS should assess the occurrence of a SIRS event with a correct timing (e.g. SIRS event is identified within the correct hour respectively shift). For example, this fifth case prevents that alert firings on day 30 of the intensive care stay will be evaluated as *true positives* if the gold standard reports a SIRS episode on day 2. Here, the CDSS did not identify the SIRS episode within the correct timing. Thus, this case is used for the determination of both *false positives* (day 30, contributing to specificity) and *false negatives* (day 2, contributing to sensitivity). Hence, the fifth case (false positive and false negative) can be defined as follows: *the gold standard reports at least one SIRS episode, and the CDSS detect SIRS episodes but (at least one) not within the same hour respectively shift*. All other cases are defined as usual (e.g. false positive: *the gold standard reports no SIRS episode but the CDSS detects one or multiple SIRS episodes*).

Based on the different cases, the sensitivity and the specificity will be determined independently. For sample size calculation, the results of the proof-of-concept study were used as a basis (sensitivity 1.00 and specificity 0.94, when calculating on the level of days). For this study with the modified statistical analysis approach, a sensitivity of 90% (alternative hypothesis: 0.98, null hypothesis: 0.90) and a specificity of 80% (alternative hypothesis: 0.90, null hypothesis: 0.80) were chosen as a clinically relevant diagnostic accuracy, with a given accuracy of estimate of 95% (type I error = 0.05) and a power of 90% (chi square test). Consequently, 97 patients suffering from at least one SIRS episode (for the estimation of sensitivity) as well as 137 patients with or without SIRS episodes are required (for the

estimation of specificity). Based on the expected incidence and prevalence, at least 300 patients need to be considered.

Timeline

Before study start, the clinicians were introduced in their tasks. No interventions, treatments or other care-related actions are prohibited and patients are treated with standard procedures (including data collection and measurements). Personal briefings on the routine documentation were carried out during this *pilot phase* (July 1, 2018, duration: 1 month, see Figure 1). A designated physician will present the study to the patients, their parents or their legal guardians and ask for consent within the *recruiting phase* (August 1, 2018, estimated duration: 10 months). Simultaneously, clinicians reported their findings during their working shift per patient (routine assessments, *diagnostic approach II*). The clinicians do not perform extensive analyses of documentations or reported data (*assessment phase I*, August 1, 2018, estimated duration: 10 months). In the routine assessments it is documented whether the patient suffered from SIRS, sepsis or organ dysfunction (via digital documentation form, see Figure 2). The first two weeks of this phase will be treated as test phase.

Later on, two experienced clinicians started with their weekly, extensive, blinded review and the definition of gold standard assessments per patient and per hour (*assessment phase II*, May 1, 2019, estimated duration: 2 months). As soon as 97 patients suffering from one or more SIRS episodes as well as 137 patients with one or without SIRS episodes have been identified, the recruitment will be terminated. Simultaneously, the data sets from all recruited patients will be integrated into a data repository to make them accessible for the CDSS. The CDSS assessments per patients and per hour will start (*diagnostic approach I*). In the final *analysis phase* (July 1, 2019, estimated duration: 2 months), the diagnostic accuracy of the CDSS will be evaluated by comparing the assessments to the gold standard decisions from the experts (primary goal of the study). Additionally, the diagnostic accuracy of the routine assessments will be determined by comparing them to the same gold standard decisions. The different accuracies can be compared (secondary goal of the study). Finally, the study results will be communicated via publication in a peer-reviewed journal.

Recruitment and consent

Eligible patients, their parents or their legal guardians will receive an information letter and a consent form (available in German, English, Turkish and Arabic) during a personal discussion with a physician. Additional information sheets for younger patients are available, one for children aged six to eleven and one for children aged 12 to 18. The families will receive privacy statement forms (data protection, accessibility and confidentiality; see additional file 2).

Patient and public involvement

There were no funds or time allocated for patient and public involvement so we were unable to involve patients. We intend to disseminate the results to the participants and will invite patients to help us developing an appropriate method of dissemination.

Data management and collection

The CDSS is an application with interfaces to a data repository, which is based on an semantic interoperability standard for clinical information representation (*openEHR* [28]). For more information, we refer to [17]. For the routine assessment (assessment phase I, diagnostic approach II), we created a documentation form, which is based on the same interoperability standard and the same interfaces to the data repository (see Figure 2). Thereby, all results (gold standard, diagnostic approach I "CDSS", diagnostic approach II "routine assessments") will be available in the same format. Patient data (identification, birthdate), intensive care parameters (heart rate, respiratory rate, body temperature, leukocytes, neutrophils, mechanical ventilation, cooling devices, pacemaker), alert parameters (SIRS, sepsis, organ dysfunction with duration, beginning and end of the episode and shift), and general documentations of the patient conditions, events or unintended effects will be documented and processed.

Data monitoring and auditing

Quality assurance measures are continuously carried out. Plausibility checks will be executed while integrating data into our repository (e.g. simple counts to can uncover whether data from the primary source is missing in the repository). Furthermore, the data set will be reviewed by physicians with respect to randomly selected observations to guarantee the

plausibility from a clinical perspective. By following the openEHR standard, we are able to automatically execute validation checks to uncover missing or wrong values when integrating the data sets or filling out the documentation form (e.g. definition of ranges for specific values, or double data entries). The study procedures will be monitored by the authors as well as by designated physicians and nurses. They will supervise the adherence to the study protocol, the procedures for routine documentation and data integration, data quality and privacy.

Data protection: data access and confidentiality

We designed a data protection concept in cooperation with the local data security officer. The concept defines pseudonymization and data access procedures, outlines the patient consent, and explains technical security mechanisms. All data sets collected or created as part of this study are treated as strictly confidential. The data sets will be stored pseudonymized and in secure conditions in the data repository located in the network of the Hannover Medical School. To prevent unauthorized disclosure of patient information, it is only accessible for the physicians and employees in charge of this study. Collected data from patients withdrawing their consent (drop outs), will be completely deleted from the data repository. All study files, the final study data sets as well as the study results will be archived for ten years in an approved long-term repository and in accordance to the relevant legal and statutory requirements. The patient will be informed about these procedures as well as their rights (including the possibility to withdraw the consent and to obtain information about collected data sets at any time), and will be asked to consent to these.

Ethics and dissemination

Ethics approval was given by the Ethics Committee of Hannover Medical School (approval number 7804_BO_S_2018). The trial is registered with ClinicalTrials.gov (NCT03661450). Protocol modifications will require a formal amendment to the protocol which will be reported to the Hannover Medical School Ethics Committee for approval. All aspects are designed according to the General Data Protection Regulation from the European Union (2016/679) and are accepted by the data security officer of the Hannover Medical School. Further details on data protection aspects can be requested from the authors. Consent to participate will be given by the patients, by their parents or by their legal guardians by signing

a study consent form. Results of the main study will be communicated via publication in a peer-reviewed journal. We intend to disseminate the results to the participants through an appropriate method of dissemination to be defined.

DISCUSSION

To be used in the long run, CDSSs have to deliver relevant information in a timely manner and at an adequate frequency. Current approaches for evaluating the usefulness of CDSS indeed present positive results. However, due to a restricted study design not designed towards daily work conditions, results may not represent the system feasibility in routine clinical care. With our work, we contribute a modified study design for evaluating the diagnostic accuracy of a CDSS with a strong focus on routine clinical care. We hypothesize that such an evaluation will demonstrate the potentials of CDSS use in routine clinical care. In case of a positive study outcome, we will be able to reason that our CDSS is not only feasible from a technical but also from a clinical perspective as it supports clinicians in critical diagnostic decision-making. For evaluation, a so-called gold standard representing the true state of the patient is required. However, in complex, knowledge-and experience-based contexts as diagnostic decision-making, reproducible, objective and quantitative "gold standards" are rare. We use an excessive evaluation of the patient data by two experience clinicians as benchmark. To reduce possible biases, the clinicians are blinded to each other and to the CDSS. In situations of disagreement, a third clinician will be consulted, decisions will be revealed and a consensus decision will be reported. We are aware that our approach is time-consuming requiring highly engaged clinicians. Because of the stressful PICU environment, assessments may be delayed, and thus, the study timeline may not be adhered. For an early recognition of issues and study monitoring at the ward, an assistant physician is in charge. Also the routine assessments of clinicians have to be managed as they are at the same risk to be biased. Clinical documentation might be handled more meticulous at the beginning and more careless in the end of the study. To prevent the latter, the new documentation form was designed in cooperation with the users and integrated in the PDMS used daily. Together with the designated study monitor, this integration raised the satisfactory and the utilization rate of the form as well as the adherence to the study protocol.

For sample size calculation, it might be possible that the incidence was overestimated, so that in our settings more than 300 patients are needed to reach 97 patients suffering from at least

one SIRS episode. The recruitment will be continuously aligned towards the number of recruited SIRS patients to be able to stop the recruitment as soon as the required number has been reached. Our expected values for sensitivity and specificity are rather conservative because we decided to primarily use an equally conservative statistical analysis approach. However, the expected values are treated as acceptable in clinical routine as the diagnostic accuracy of the system will be over a critical minimum (and with respect to the aspired second goal of our study, even better than in clinical routine decision-making). At the same time, alert fatigue will be prevented because specificity is equally high. We are confident that our thoughts meet the need for an optimum balance between sensitivity and specificity, e.g. as reported by Coleman et al. [23]. Nevertheless, we will enhance our results with a more liberal analysis on the level of days.

Our study has been successfully started with recruitment according to this design and promises valuable results. When reaching a good diagnostic accuracy compared to the gold standard as well as advantages over the diagnostic accuracy of routine assessments, we are optimistic that our users are willing to trust and use the system in future. Moreover, this will allow the conduction of future studies as for example the evaluation of patient outcomes, user acceptability, or real-time performance of the system.

STATEMENTS

Author Contributions

AW was responsible for design and implementation of the presented clinical decision-support system and the outline of the study protocol, and has drafted the manuscript. TJ provided clinical expertise for the use case and the design of the underlying knowledge model, leaded the proof-of-concept study and co-drafted the manuscript. AK was primarily responsible for the design of the statistical analysis, the sample size calculation and the authoring of the corresponding sections. BS and SM helped in the conception of the general study approach but especially for definition of goals and outcome measures, timing and patient recruitment; SM is responsible for patient recruitment and monitors the study at the ward. PB and MM provided clinical expertise for study design, revised the manuscript critically, and gave subject-specific advices as well as the final approval of the manuscript version to be published. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Funding

No funding to declare.

Availability of data and material

The datasets generated and/or analyzed during the current study are not publicly available due to data privacy and security matters of patients but are available from the corresponding author on reasonable request.

Acknowledgements

We want to thank our colleagues from the ZIMt department for Educational and Scientific IT Systems of the Hannover Medical School for supporting in matters of data access and integration.

REFERENCES

1 American College of Chest Physicians/Society of Critical Care Medicine Consensus Conference: definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. *Crit Care Med* 1992;20(6):864–74. PubMed PMID: 1597042.

2 Goldstein B, Giroir B, Randolph A. International pediatric sepsis consensus conference: Definitions for sepsis and organ dysfunction in pediatrics*. *Pediatr Crit Care Med* 2005;6(1):2–8. doi: 10.1097/01.PCC.0000149131.72248.E6.

3 Wolfler A, Silvani P, Musicco M, et al. Incidence of and mortality due to sepsis, severe sepsis and septic shock in Italian Pediatric Intensive Care Units: A prospective national survey. *Intensive Care Med* 2008;34(9):1690–97. doi: 10.1007/s00134-008-1148-y.

4 Cvetkovic M, Lutman D, Ramnarayan P, et al. Timing of death in children referred for intensive care with severe sepsis: Implications for interventional studies. *Pediatric Critical Care Medicine* 2015;16(5):410–17. doi: 10.1097/PCC.0000000000000385.

5 Creedon JK, Vargas S, Asaro LA, et al. Timing of Antibiotic Administration in Pediatric Sepsis. *Pediatr Emerg Care* 2018. doi: 10.1097/PEC.0000000000001663.

6 Kortz TB, Sawe HR, Murray B, et al. Clinical Presentation and Outcomes among Children with Sepsis Presenting to a Public Tertiary Hospital in Tanzania. *Front Pediatr* 2017;5:278. doi: 10.3389/fped.2017.00278.

7 Boehne M, Sasse M, Karch A, et al. Systemic inflammatory response syndrome after pediatric congenital heart surgery: Incidence, risk factors, and clinical outcome. *J Card Surg* 2017;32(2):116–25. doi: 10.1111/jocs.12879.

8 Ganjoo S, Ahmad K, Qureshi UA, et al. Clinical Epidemiology of SIRS and Sepsis in Newly Admitted Children. *Indian J Pediatr* 2015;82(8):698–702. doi: 10.1007/s12098-014-1618-x.

9 Han YY, Carcillo JA, Dragotta MA, et al. Early reversal of pediatric-neonatal septic shock by community physicians is associated with improved outcome. *Pediatrics* 2003;112(4):793–99. doi: 10.1542/peds.112.4.793.

10 Paul R. Recognition, Diagnostics, and Management of Pediatric Severe Sepsis and Septic Shock in the Emergency Department. *Pediatr Clin North Am* 2018;65(6):1107–18. doi: 10.1016/j.pcl.2018.07.012.

11 Sackett DL. Evidence-based medicine. *Semin Perinatol* 1997;21(1):3–5. doi: 10.1016/S0146-0005(97)80013-4.

12 Zavala AM, Day GE, Plummer D, et al. Decision-making under pressure: Medical errors in uncertain and dynamic environments. *Aust Health Rev* 2017. doi: 10.1071/AH16088.

13 Lighthall GK, Vazquez-Guillamet C. Understanding Decision Making in Critical Care. *Clin Med Res* 2015;13(3-4):156–68. doi: 10.3121/cmr.2015.1289.

14 Castaneda C, Nalley K, Mannion C, et al. Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine. *J Clin Bioinforma* 2015;5:4. doi: 10.1186/s13336-015-0019-3.

15 Lincoln P, Manning M-J, Hamilton S, et al. A pediatric critical care practice group: use of expertise and evidence-based practice in identifying and establishing "best" practice. *Crit Care Nurse* 2013;33(2):85–87. doi: 10.4037/ccn2013740.

16 Shortliffe EH. Biomedical Informatics: Defining the Science and Its Role in Health Professional Education. In: Hutchison D, Kanade T, Kittler J, et al., eds. Information Quality in e-Health. Berlin, Heidelberg: Springer Berlin Heidelberg 2011:711–14.

17 Wulff A, Haarbrandt B, Tute E, et al. An interoperable clinical decision-support system for early detection of SIRS in pediatric intensive care using openEHR. *Artif Intell Med* 2018;89:10–23. doi: 10.1016/j.artmed.2018.04.012.

18 Wulff A, Marschollek M. Learning Healthcare Systems in Pediatrics: Cross-Institutional and Data-Driven Decision-Support for Intensive Care Environments (CADDIE). *Stud Health Technol Inform* 2018;251:109–12. doi:10.3233/978-1-61499-880-8-109

19 Faisal M, Scally A, Richardson D, et al. Development and External Validation of an Automated Computer-Aided Risk Score for Predicting Sepsis in Emergency Medical Admissions Using the Patient's First Electronically Recorded Vital Signs and Blood Test Results. *Crit Care Med* 2018;46(4):612–18. doi: 10.1097/CCM.0000000000002967.

20 Austrian JS, Jamin CT, Doty GR, et al. Impact of an emergency department electronic sepsis surveillance system on patient mortality and length of stay. *J Am Med Inform Assoc* 2018;25(5):523–29. doi: 10.1093/jamia/ocx072.

21 Mao Q, Jay M, Hoffman JL, et al. Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department, general ward and ICU. *BMJ Open* 2018;8(1):e017833. doi: 10.1136/bmjopen-2017-017833.

22 Kilsdonk E, Peute LW, Jaspers MW. Factors influencing implementation success of guideline-based clinical decision support systems: A systematic review and gaps analysis. *Int J Med Inform* 2017;98:56–64. doi: 10.1016/j.ijmedinf.2016.12.001.

23 Coleman JJ, van der Sijs H, Haefeli WE, et al. On the alert: Future priorities for alerts in clinical decision support for computerized physician order entry identified from a European workshop. *BMC Med Inform Decis Mak* 2013;13:111. doi: 10.1186/1472-6947-13-111.

24 Liberati EG, Ruggiero F, Galuppo L, et al. What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. *Implement Sci* 2017;12(1):113. doi: 10.1186/s13012-017-0644-2.

25 Jack T, Boehne M, Brent BE, et al. In-line filtration reduces severe complications and length of stay on pediatric intensive care unit: A prospective, randomized, controlled trial. *Intensive Care Med* 2012;38(6):1008–16. doi: 10.1007/s00134-012-2539-7.

26 Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: Explanation and elaboration. *BMJ Open* 2016;6(11):e012799. doi: 10.1136/bmjopen-2016-012799.

27 Chan A-W, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: Guidance for protocols of clinical trials. *BMJ* 2013;346:e7586. doi: 10.1136/bmj.e7586.

28 Thomas Beale. Archetypes: Constraint-based Domain Models for Future-proof Information Systems. In: Eleventh OOPSLA workshop on behavioral semantics: serving the customer. Seattle, Washington, Boston: Northeastern University 2002:16–32.

Figure 1: Time schedule and study episodes for CADDIE2 trial

Figure 2: Digital documentation form (based on openEHR data repository)

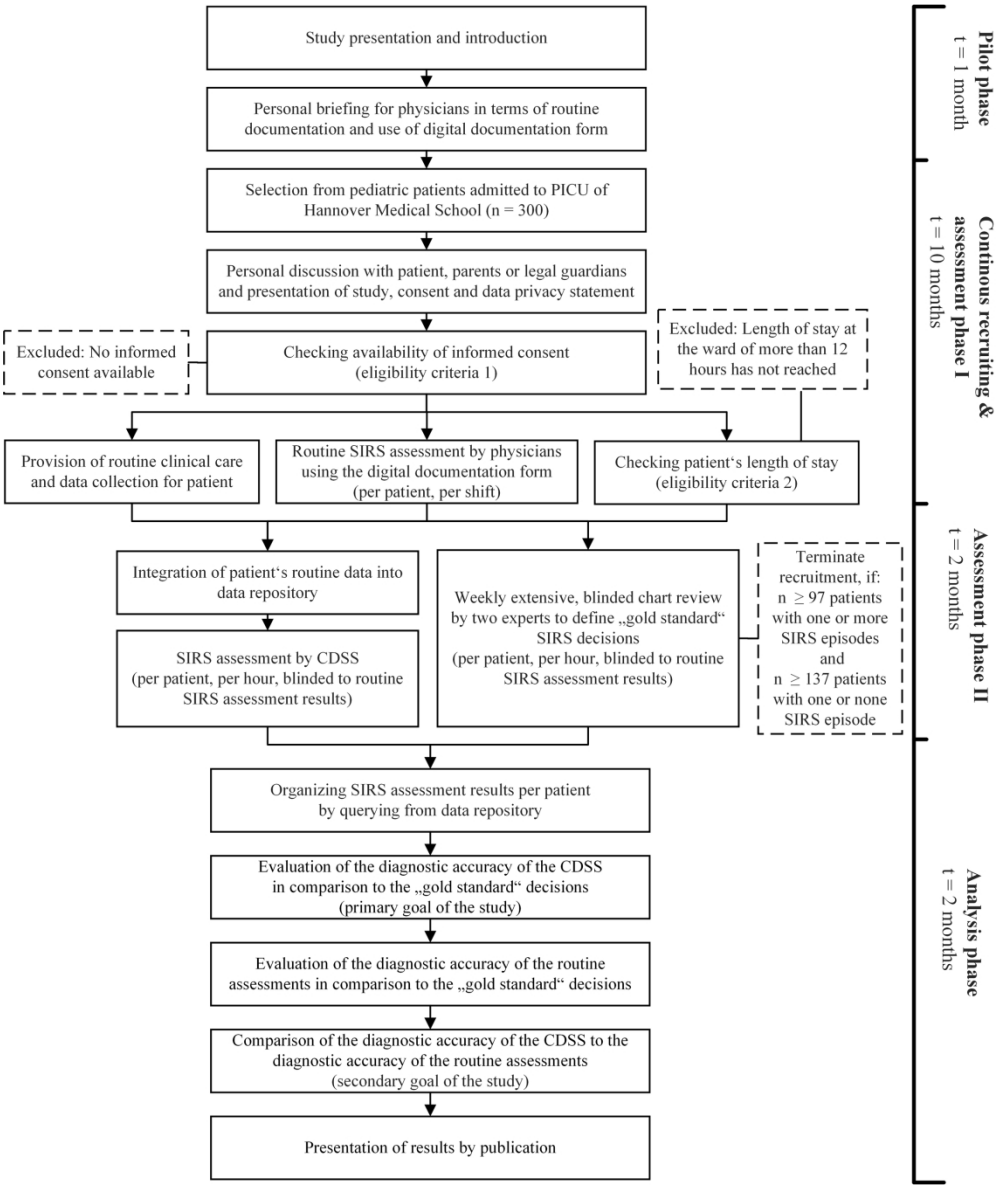

Figure 1: Time schedule and study episodes for CADDIE2 trial

208x245mm (300 x 300 DPI)

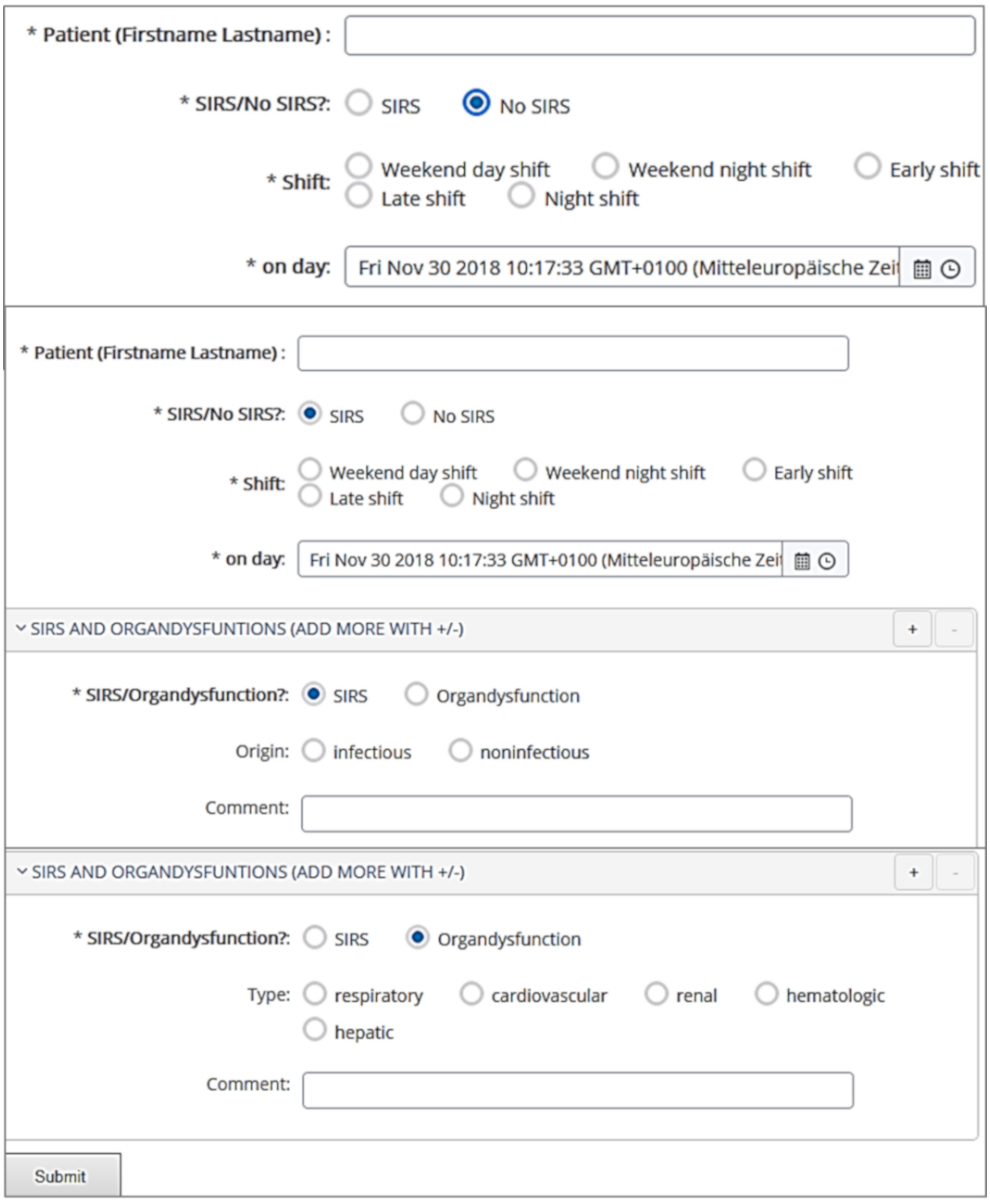

Figure 2: Digital documentation form (based on openEHR data repository)

90x109mm (300 x 300 DPI)

SPIRIT
STANDARD PROTOCOL ITEMS: RECOMMENDATIONS FOR INTERVENTIONAL TRIALS

SPIRIT 2013 Checklist: Recommended items to address in a clinical trial protocol and related documents*

Section/item	Item No	Description	Addressed on page number
Administrative information			
Title	1	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	1
Trial registration	2a	Trial identifier and registry name. If not yet registered, name of intended registry	3
	2b	All items from the World Health Organization Trial Registration Data Set	NA
Protocol version	3	Date and version identifier	3
Funding	4	Sources and types of financial, material, and other support	18
Roles and responsibilities	5a	Names, affiliations, and roles of protocol contributors	1,18
	5b	Name and contact information for the trial sponsor	NA
	5c	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	18
	5d	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	9,15,16

Introduction			
Background and rationale	6a	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	_____5,6,7__
	6b	Explanation for choice of comparators	_____5,6,7,8__
Objectives	7	Specific objectives or hypotheses	_____7,8_____
Trial design	8	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, noninferiority, exploratory)	_____8_____
Methods: Participants, interventions, and outcomes			
Study setting	9	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will be collected. Reference to where list of study sites can be obtained	_____8_____
Eligibility criteria	10	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	_____9_____
Interventions	11a	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	NA (no interventions, diagnostic study), alternative: _____8_____
	11b	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving/worsening disease)	_____NA (each patient will be assessed by both diagnostic approaches), alternative: _____8_____

	11c	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return, laboratory tests)	_____NA (no intervention, routine care), alternative: _____16____
	11d	Relevant concomitant care and interventions that are permitted or prohibited during the trial	_____10____
Outcomes	12	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	_____9,10___
Participant timeline	13	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	_____10,11,12_
Sample size	14	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	_____9,10____
Recruitment	15	Strategies for achieving adequate participant enrolment to reach target sample size	_____12,13__
Methods: Assignment of interventions (for controlled trials)			
Allocation:			
Sequence generation	16a	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	_____NA____
Allocation concealment mechanism	16b	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	_____NA____
Implementation	16c	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	_____NA____

Blinding (masking)	17a	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	7,11,16
	17b	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	16
Methods: Data collection, management, and analysis			
Data collection methods	18a	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known. Reference to where data collection forms can be found, if not in the protocol	13,14
	18b	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	NA (no interventions)
Data management	19	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	13,14, 15
Statistical methods	20a	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	9,10
	20b	Methods for any additional analyses (eg, subgroup and adjusted analyses)	NA
	20c	Definition of analysis population relating to protocol non-adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	NA
Methods: Monitoring			
Data monitoring	21a	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	15
	21b	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	NA

Harms	22	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events and other unintended effects of trial interventions or trial conduct	15
Auditing	23	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	15,16
Ethics and dissemination			
Research ethics approval	24	Plans for seeking research ethics committee/institutional review board (REC/IRB) approval	16, 18, 19, Additional file 3
Protocol amendments	25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC/IRBs, trial participants, trial registries, journals, regulators)	19
Consent or assent	26a	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	12,13
	26b	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	NA
Confidentiality	27	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	13,15
Declaration of interests	28	Financial and other competing interests for principal investigators for the overall trial and each study site	18
Access to data	29	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	18
Ancillary and post-trial care	30	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	NA
Dissemination policy	31a	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	11,16

	31b	Authorship eligibility guidelines and any intended use of professional writers	_____18____
	31c	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	_____NA___
Appendices			
Informed consent materials	32	Model consent form and other related documentation given to participants and authorised surrogates	_____Additional file 2__
Biological specimens	33	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	_____NA____

*It is strongly recommended that this checklist be read in conjunction with the SPIRIT 2013 Explanation & Elaboration for important clarification on the items. Amendments to the protocol should be tracked and dated. The SPIRIT checklist is copyrighted by the SPIRIT Group under the Creative Commons "Attribution-NonCommercial-NoDerivs 3.0 Unported" license.

Patient Information (Parents)

Dear parents,

we would like to include your child in a clinical trial, you will find the details on the following pages. Please read them carefully. If you have any questions please do not hesitate to ask the doctor in charge.

Information about the clinical trial which is named:

Evaluation of the diagnostic accuracy of a Clinical Decision-Support System (CDSS) to support recognition of SIRS and Sepsis in paediatric intensive care patients in comparison to medical specialists.

Your child was or will be admitted on the paediatric intensive care unit (PICU) at Medizinische Hochschule Hannover. Due to neccessary monitoring and therapy we will gather laboratory results and vital signs continueally.

During the clinical trial these laboratory and monitoring details will be analyzed with regard to a pronounced, generalized inflammatory reaction (Systemic inflammatory response syndrome (SIRS)) using the support of a computer-based system. The objective is to support the doctors in early recognition of SIRS.

The recognition of SIRS on a PICU is a complex task that is quite suitable for a Clinical Decision-Support System. SIRS is caused by different reasons, for example surgery, injurys or infection. It is a typical occurence at PICU-patients and puts the patient under additional strain independent from the original reason for admittance. For diagnosing SIRS in paediatric patients correctly we have to find abnormalities in different vital signs and laboratory results which have age related reference values, making it an advanced task.

The development and clinical trial of a CDSS for SIRS in paediatric patients is the result of a cooperation between the Department of Paediatric Cardiology and Intensive Care Medicine of Medizinische Hochschule Hannover and the Peter L. Reichertz Institut für

Medizinsche Informatik der Technischen Universität Braunschweig und der Medizinischen Hochschule Hannover (PLRI).

Are there any risks by taking part in the clinical trial?

The CDSS will only analyze data that is taken and saved during your child's admittance on the PICU anyway. There won't be any additional blood tests, monitoring or examinations. These standardized raised data will be saved in a pseudonymized way that ensures the patient's privacy.

What happens if I don't agree to take part in the clinical trial?

The whole team will respect your decision if you do not want your child to take part in the clinical trial. This will not be evaluated or have any negative effect on your child's treatment. You can also revoke your agreement anytime.

Privacy Policy

The data raised and saved for the trial are pseudonymized and it is impossible to reconnect them to an individual. The data will be saved for 10 years and deleted afterwards. The issues of the Bundesdatenschutzgesetz are considered in every respect.

I received the information and consent to the clinical trial mentioned above.

☐Yes ☐No

_____

(Date and signature of the parent/legal guardian)

_____

(Date and signature of the medical doctor)

Patient Information (Patient)

Dear patient,

we would like to include you in a clinical trial, you will find the details on the following pages. Please read them carefully. If you have any questions please do not hesitate to ask the doctor in charge.

Information about the clinical trial which is named:

Evaluation of the diagnostic accuracy of a Clinical Decision-Support System (CDSS) to support recognition of SIRS and Sepsis in paediatric intensive care patients in comparison to medical specialists.

You were or will be admitted on the paediatric intensive care unit (PICU) at Medizinische Hochschule Hannover. To treat you best we will monitor your vital signs as heartbeat, blood pressure and respiratory rate and take blood tests to monitor your white blood count and other inflammatory parameters. The results will be saved.

We know about the occurence of extreme inflammatory reactions in paediatric patients. These can be cause by several reasons as surgery, injury or infection. We can measure these by changes in your vital signs, temperature or white blood count.
With a clinical trial something new is tested, for example some treatment. For this clinical trial your monitoring or vital sign data will be analyzed by a computer programme to support your doctors to detect changes that might be caused by an inflammatory reaction as soon as possible. The sooner we are aware we can adjust your treatment.

Are there any risks by taking part in the clinical trial?

The CDSS will only analyze data that is taken and saved during your admittance on the PICU anyway. There won't be any additional blood tests, monitoring or examinations. These standardized raised data will be saved in a pseudonymized way that ensures the patient's privacy.

What happens if I don't agree to take part in the clinical trial?

The whole team will respect your decision if you do not want to take part in the clinical trial. This will not be evaluated or have any negative effect on your treatment. You can also revoke your agreement anytime.

Privacy Policy

The data raised and saved for the trial are pseudonymized and it is impossible to reconnect them to an individual. The data will be saved for 10 years and deleted afterwards. The issues of the Bundesdatenschutzgesetz are considered in every respect.

I received the information and consent to the clinical trial mentioned above.

☐Yes ☐No

_____

(Date and signature of the patient)

_____

(Date and signature of the medical doctor)