Electronic clinical decision support in acute and critical care

Electronic Clinical Decision Support in Acute and Critical Care

Investigating the closing loop

ASHLEY DE BIE

Electronic Clinical Decision Support in Acute and Critical Care

Investigating the closing loop

ASHLEY DE BIE

COLOFON

Title:	Electronic Clinical Decision Support in Acute and Critical Care
Author:	Ashley De Bie
Cover design:	James Jardine | www.jamesjardine.nl
Layout:	James Jardine | www.jamesjardine.nl
Print:	Ridderprint | www.ridderprint.nl
ISBN:	978-94-6416-257-8

This work was performed within the framework of the Eindhoven MedTech Innovation Centre (e/MTIC or IMPULS II) in collaboration with the Eindhoven University of Technology, Catharina Hospital Eindhoven and Philips Research (Eindhoven).

Electronic Clinical Decision Support in Acute and Critical Care

Investigating the closing loop

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de rector magnifcus prof. dr ir F.P.T. Baaijens, voor een commissie aangewezen door het College voor Promoties, in het openbaar te verdedigen op donderdag 19 november 2020 om 16:00 uur

door

Ashley Jacobus Raphaël De Bie
geboren te Eindhoven, Nederland

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

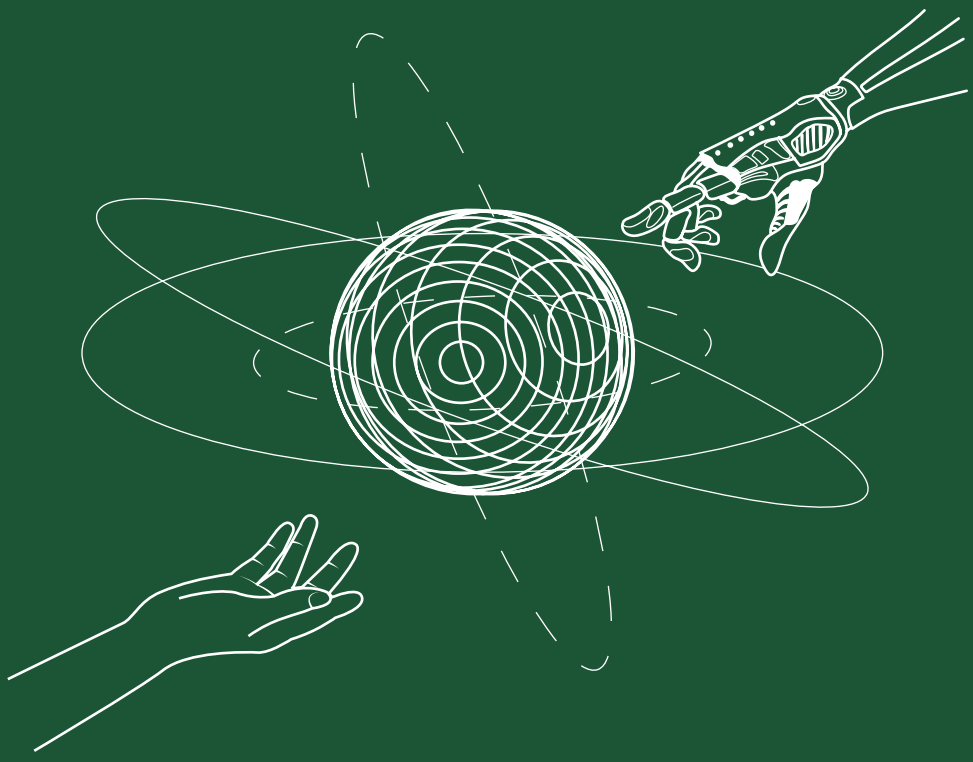
Voorzitter:	Prof. dr ir A.C.P.M. Backx	
1e promotor:	Prof. dr H.H.M. Korsten	
1e copromotor:	Dr A.J.G.H. Bindels	Catharina Ziekenhuis Eindhoven
2e copromotor:	Dr R.A. Bouwman	Catharina Ziekenhuis Eindhoven
3e copromotor:	Dr P.M.E. Van Gorp	
Leden:	Prof. dr N.J.M. van der Meer	TIAS School for Business and Society
	Prof. dr C.J. Kalkman	Universitair Medisch Centrum Utrecht
	Prof. dr N.H.J. Pijls	
Adviseurs:	Prof. dr X. Lu	Zhejiang University
	Prof. dr H. Duan	Zhejiang University

Het onderzoek of ontwerp dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

't Is goed in 't eigen hert te kijken
En zóó z'n oogen toe te doen

Voor mammy & pappy
voor mijn moeder

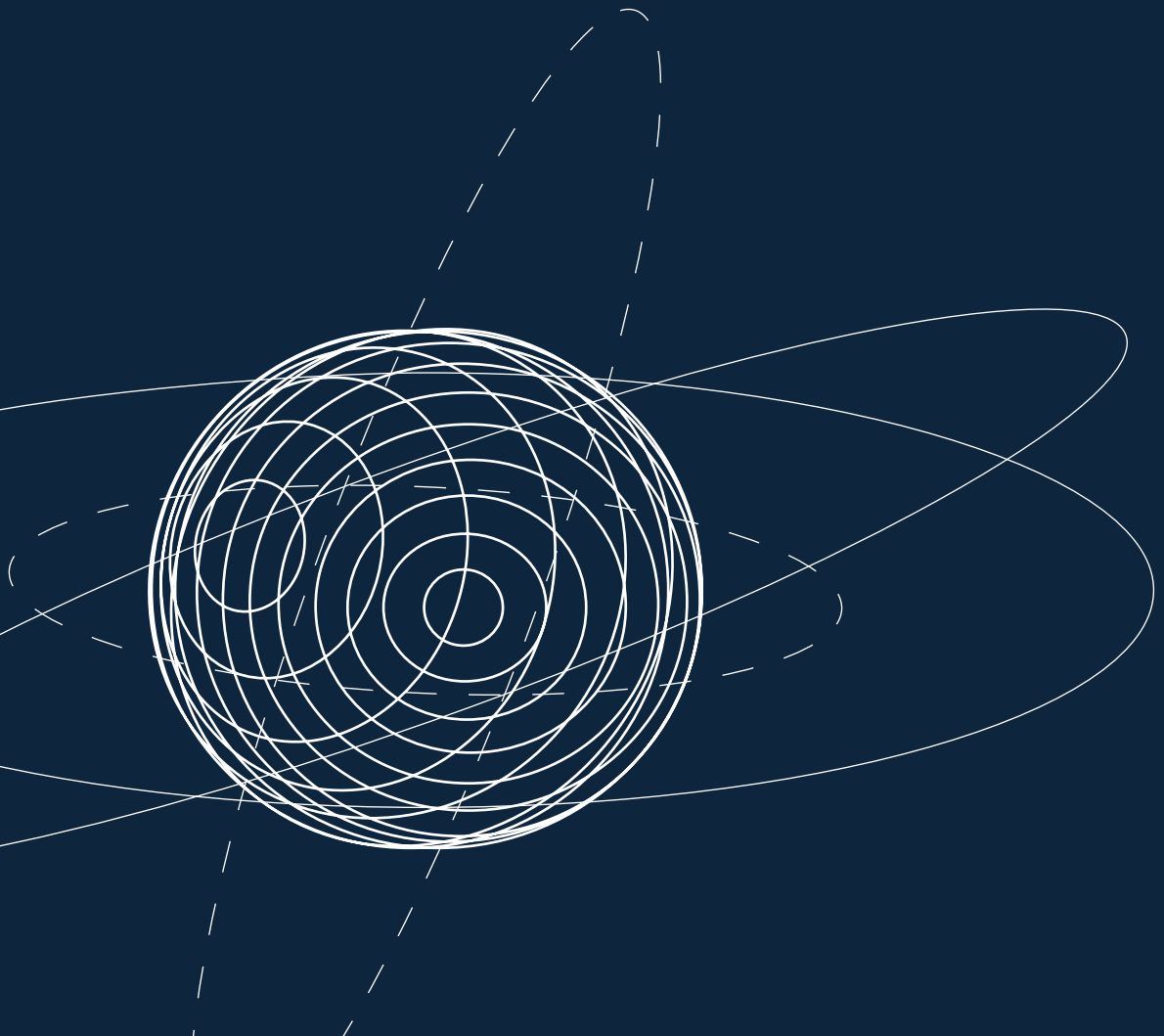
Table of Contents



PART 1

Introduction

1.1



General introduction

Clinical decision support (CDS) in healthcare is a strategy for assisting clinicians in their complex decision-making processes. Most health-related decisions and medical procedures are still based on human memory, but the human brain is getting overwhelmed by the increasing amount, production speed, and multidimensionality of available medical data.[1, 2] Moreover, discrepancies between health care demand and supply are becoming more common due to an aging population with multi-morbidity and increasing severity of illness. The growing population of these severely ill patients increases the complexity of healthcare. This complexity increases the importance of effective communication between various involved caregivers to optimize patients' health care while guaranteeing their safety.[3-6] Furthermore, implementing the rapidly advancing medical knowledge and technologies into clinical practice is extremely challenging. These drawbacks combined with the need to temper costs in healthcare are placing a huge strain on healthcare professionals and may adversely affect the quality of their health-related decisions.[5-8] Auxiliary strategies, such as clinical decision support, are therefore warranted to support healthcare professionals in delivering value-based personalized healthcare.

CLINICAL DECISION SUPPORT

Clinical decision support aims to enhance health-related decision making at the point of care based on the integration of electronically stored knowledge from guidelines and protocols with specific patient information from various sources of personalized health information, such as an Electronic Medical Record.[9-13] The intention of this process is to improve patients' outcomes, prevent errors or adverse events, and help teams of healthcare professionals to be more effective. There are basically three types of decision-support functions for CDS tools that can overlap:[14, 15]

1. Information-management tools that store and grant access to medical data and information in a structured way, such as an Electronic Medical Record.
2. Tools for focusing attention, such as tools that flag abnormal values or possible drug interactions.
3. Tools that provide patient-tailored diagnostic assistance or treatment advice.

Numerous CDS tools of varying complexity are available within the entire field of medicine. They are commonly characterized based on the underlying decision-making process model (*Figure 1*).[15] They range from simple tools that do not require any form of human-technology interaction, such as clinical guidelines or checklists (*Section 0 in Figure 1*), to fully automated CDS systems (*Section 4 in Figure 1*).

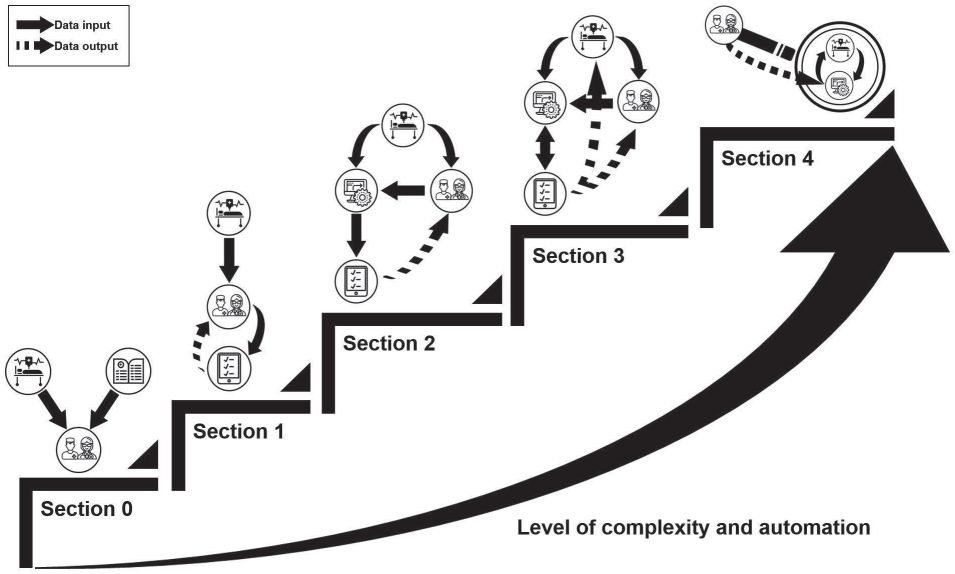


Figure 1. Various Sections representing the level of complexity and automation of a CDS tool
Section 0 CDS tool without human-technology interaction (paper checklist or guideline)
Section 1 A straightforward computerized CDS system without automation.
Section 2 A CDS system that is based on a simple model which can generate some data automatically
Section 3 A CDS system based on a complex transparent model with an automated, closed-loop, functionality
Section 4 A fully closed-loop automated CDS system

Guidelines and checklists

A medical guideline is an international, national, or local consensus statement of the best practice to diagnose, treat and manage certain clinical conditions and needs.[14, 16] These documents represent an examination of the current evidence combined with a medical expert consensus to support clinicians and patients in (individual and shared) decision making.[14, 16, 17] The benefit of medical guidelines is their ability to enhance standardization in medical care.[14, 18] In view of that benefit, a huge diversity of medical guidelines and protocols have been introduced, with which policy makers urge caregivers to comply as much as possible. However, the occurrence of numerous clinical conditions and the usage of guidelines as quality standards have contributed to an abundance of extensive and complex text-based guidelines in printed version or stored in electronic databases.[18-20] Even in case of electronic or online guidelines, clinicians struggle with limited time and means to implement and comply with all of

them in their routine daily practice.[14, 18, 21, 22] Thus checklists, used as benchmarks in other high-risk industries, have been introduced in medicine to support compliance to the guidelines and to ensure high quality of care.[23]

Checklists were first introduced in aviation in the 1930s. In 1935, two highly qualified pilots could not prevent the crash of the newest and most sophisticated bomber of Boeing, Model 299, during the first evaluation flight. Instead of demanding more training from their pilots, Boeing engineers developed a short list of crucial checks that had to be performed before take-off.[23, 24] Ever since checklists have been regarded to be among the most important contributors to aviation safety, and gradually they were adopted by other high-risk industries.[1, 23, 24]

The emergence of the surgical checklist illustrates the adoption of the checklist concept in healthcare. The first surgical checklist that was created by the World Health Organisation reduced mortality and surgical complications worldwide in a cost-effective approach.[23] A large multinational trial, including 7688 patients, showed that completing this checklist before surgery reduced the overall rate of postoperative complications.[25] In a subsequent study multiple checklists were implemented at predefined moments that corresponded to the different stages of care in the surgical pathway, from admission to discharge (preoperative, operative, recovery or intensive care, and postoperative).[26] The content of these checklists depended on the stage of care. This implementation of multidisciplinary checklists within the surgical pathway (Surgical Patient Safety System (SURPASS)) reduced the proportion of patients with one or more complications in Dutch hospitals with a high baseline standard of care.[26] This beneficial effect was even more pronounced among patients for whom 80% or more of the checklist items were completed. This result demonstrated that checklist compliance is an essential part of checklist effectiveness.[26-29] Concerning the ICU, checklists for morning rounds and specific interventions have been proven to improve compliance with care processes, resulting in e.g. lower incidences of secondary infections, such as pneumonia or catheter-related bloodstream infections.[30, 31]

Even though several subsequent studies were unable to reproduce these beneficial effects, these early studies provided a better understanding of the benefits and limitations of the use of checklists in health care.[26, 32-35] Multiple social-organizational barriers and checklist elements influence the acceptance and compliance of checklists, while checklists only have a favourable effect if clinicians comply with them properly.[27-29, 36] Therefore, checklists need to be well designed, easily accessible and

adequately integrated into the daily analogous and digital work flow, and they must contain patient information that is relevant for the specific health care provided.[1, 23, 37] With these elements in place, checklists can become a cognitive aid for clinicians.

Clinical decision support systems

The number of new available CDS systems in healthcare has increased substantially since the 1960s.[11, 12, 15, 38-42] The goal of a CDS system is to generate a clinical advice based on input data that supports clinical decision making at the point of care without negatively interfering with the clinicians' daily workflow.[11] The architecture of CDS systems appears similar to the iterative four-stage model of human information processing (*Figure 2*).[43]

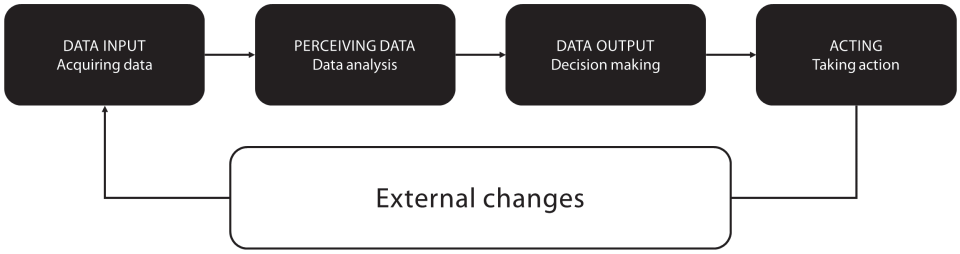


Figure 2. The model of human information processing
The model describes how humans receive, perceive, process, and use information.

In contrast to a guideline or paper checklist, all CDS systems share the common architecture of processing input data from health information systems to generate clinically relevant output, i.e. clinical advice, based on a computerized model. These models can be classified as knowledge-based or non-knowledge-based.[11, 12] Knowledge-based models apply rules (IF-THEN) that are developed and programmed by humans using literature-based, practice-based, or patient-directed evidence.[12, 44] The decision making process of non-knowledge based models leverages artificial intelligence, machine learning, or statistical pattern recognition.[11, 12, 45]

Simple CDS systems guide users at the point-of-care to choose the most appropriate action. For example, computerized static checklists stored in an interactive mobile application can guide their users in various medical settings (*Section 1 in Figure 1*).[1, 46] More advanced systems are even able to perform an action automatically within predefined boundaries set by the clinicians. The aim of this autonomous functionality is to pre-process or even take over care processes so that clinicians can focus on other tasks or cases. The CDS system's level of automation can vary and depends

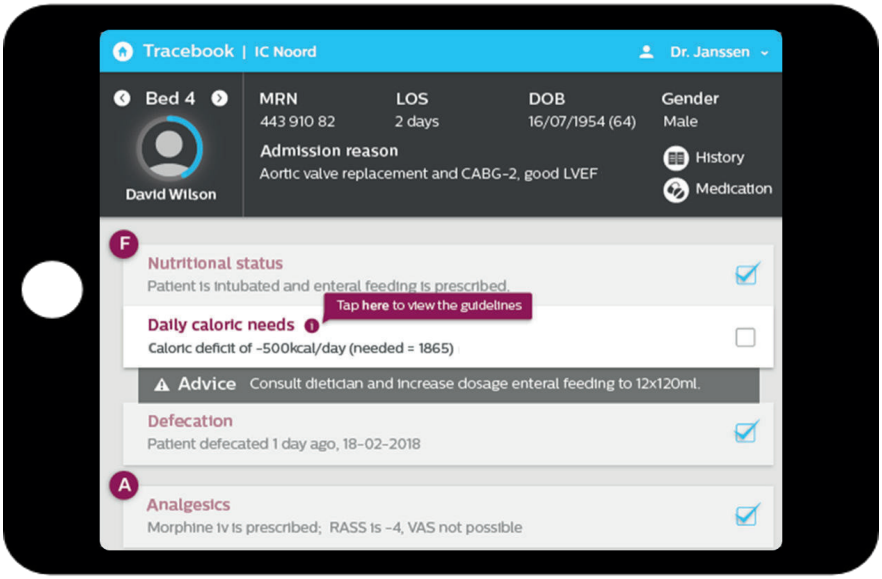
on the purpose of the CDS system. For example, a CDS system can assist healthcare professionals by automatically prompting a warning score with corresponding advice for further monitoring based on a set of vital signs that were acquired both manually and automatically (*Section 2 in Figure 1*).

As technology advances, so-called closed loop CDS systems are emerging in healthcare that can act automatically without the need for human interference. A CDS systems that contains this closed loop functionality is TraceBook, which can be activated for multiple care processes based on the preference of the local users.[47] TraceBook aims to provide a digital overview of the patient within a specific clinical pathway, depicting all healthcare professionals (as well as the patient and the patient's family) and the time relations and dependencies between them. The overview is compiled using various multidisciplinary computerized checklist containing patient specific recommendations for multiple care processes which can be completed by the involved care givers or can be processed automatically within predefined boundaries (*Section 3 in Figure 1*).[48] As a result, the content and design (*Figure 3*) of these checklists are dynamic, using patient-specific items in a process-oriented and context-aware manner. Finally, there already are some commercially available CDS systems that can fully take over a certain care process, such as an automated ventilation mode which automatically applies the safest ventilator settings in each breath while the clinician at the bedside monitors its actions (*Section 4 in Figure 1*).[49-53]

Challenges for clinical decision support tools

A major impediment to the introduction of CDS systems in healthcare is the requirement to integrate the CDS systems into daily routines.[54] The advice or interventions suggested by the CDS systems must be logical to the users and should not negatively interfere with their daily workflow, especially if certain clinical tasks are accomplished automatically by the CDS systems.[1, 11, 55, 56] The complexity of clinical practice is reflected in various technological, design-related, clinical, and socio-organizational challenges that can impede the implementation of CDS systems into practice. These challenges are represented in *Figure 4* as a set of recommendations for a useful CDS system.[9, 10, 13, 40, 57, 58] However, these recommendations are still mainly based on experience-based recommendations and scientific evidence supporting them is required.[10, 11, 57] Scientific research should therefore focus on studying and exploring these recommendations for each CDS system in relation to its users' daily practice.

A. DO-CONFIRM format



B. READ-COMPLETE format

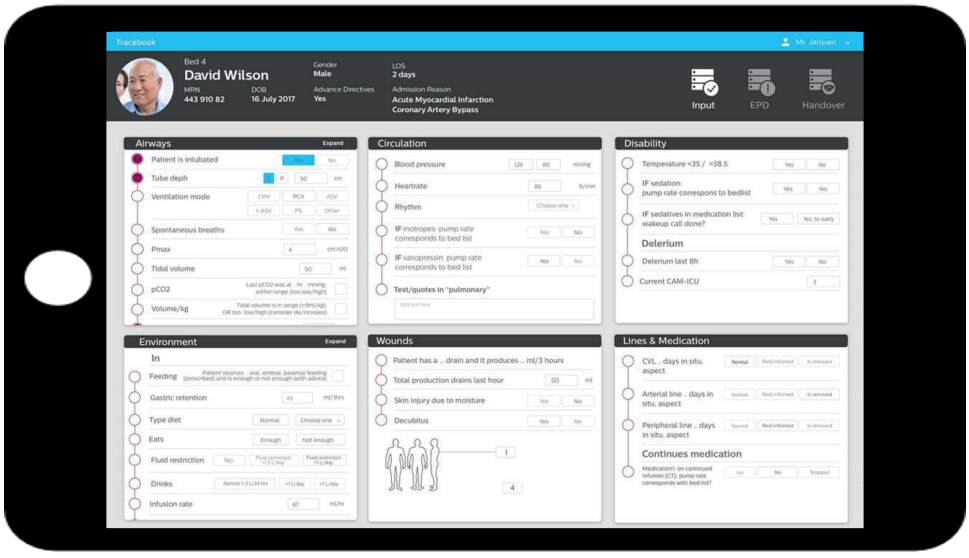


Figure 3. The dynamic functionality of TraceBook's dynamic clinical checklists

This figure demonstrates how the design and content of a computerized checklist can vary based on the user and context. The first checklist (A) is a DO-CONFIRM format in which the user, i.e. the physician, is asked to check if an item has been considered. The second checklist (B) is a READ-COMPLETE format requesting the user to provide data. Both formats can be merged into one checklist.

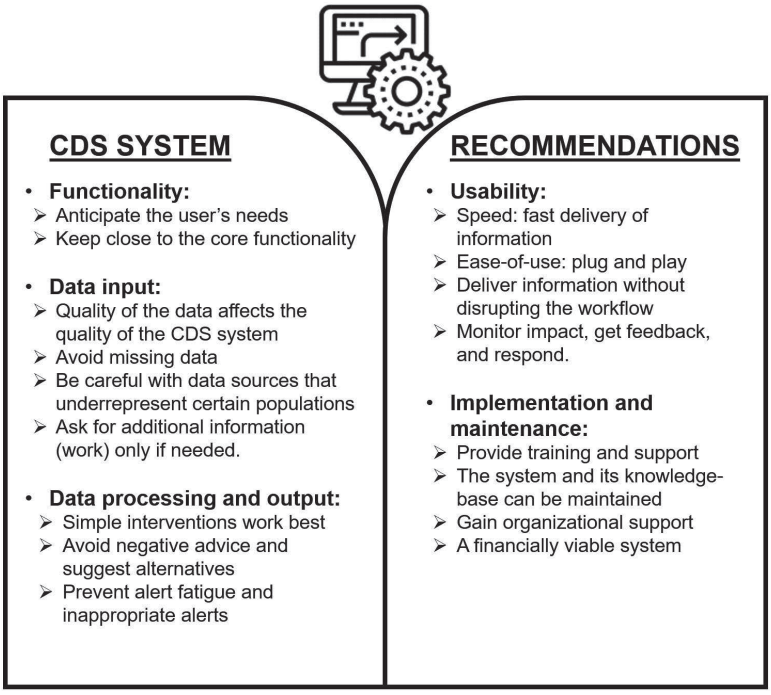


Figure 4. Experience-based recommendations for a useful CDS system

A set of experience-based recommendations for a useful CDS system was established on various technological, design-related, clinical, and socio-organizational challenges that reflect the complexity of clinical practice.[9-13, 40, 57, 58]

CLINICAL DECISION SUPPORT SYSTEMS IN ACUTE AND CRITICAL CARE

Acute and critical care

Acute care includes the timely detection and early care management of patients who become acutely ill and require stabilization for transfer to another higher dependency unit. Critical care is the specific comprehensive care of patients with life-threatening diseases that require constant monitoring in high dependency units such as an Intensive Care Unit (ICU). Patients can be admitted to the ICU from an emergency department, immediately after invasive surgery (such as cardiothoracic surgery) to be stabilized for further recovery, or from the general ward if patients rapidly deteriorate during their hospital stay. The primary aim of both acute and critical care is to support life by preventing and treating life-threatening diseases and organ failure, while preventing secondary injury due to complications.

Challenges in acute and critical care

Clinical reasoning in acute and critical care is an ongoing and dynamic process based on the recognition of disease patterns and their diversity.[59] Moreover, patients often present with a variety of symptoms that are difficult to assemble into a single diagnosis, and therefore they require a personalized approach in which multiple pathologic conditions are treated by a multidisciplinary team of healthcare professionals.[59, 60] This intrinsic complexity of the critically ill patient, which is often exacerbated by the patient's inability to express complaints, generates the need for close and continuous monitoring. This monitoring is done by various medical devices and exposes healthcare professionals to a large amount of data.[61] Healthcare professionals are stressed to make appropriate decisions with this data in a timely manner for patients with little physiologic reserves in intense and distracting environments.[59-61] Their decision-making process is not exclusively focused on the underlying disease and organ failures, but also requires attention to the patient's former physical condition and wishes, to the recognition of physiological changes (trends), and to the provision of meticulous supportive care (*Figure 5*).[2, 59] However, these challenges and other factors such as the daily variation in personnel, staff experience, work intensity, and patient acuity, hamper healthcare professionals in obtaining a clear overview of the patient's condition and needs while providing continuity of care. As a result patient harm can occur from errors of commission (an incorrect diagnosis, or inadequate treatments) and particularly from errors of omission (failure to detect the patient's deteriorating condition or to start essential treatments).[59, 60]

Opportunities for CDS systems in acute and critical care.

CDS systems are promising technologies that may help to overcome the described challenges and that can be applied to several situations in both acute and critical care. Regarding acute Care, the recognition and early targeted management of a patient's deteriorating condition depends on the subjective judgement of nurses and young doctors at the bedside on general wards. Simple systems that are often implemented to support these professionals in these stressful times are a structured ABCDE approach or track-and-trigger tools, such as an Early Warning Score.[62-68] However, the effectiveness of these systems is limited by incomplete and non-accurate registration of vital signs. In clinical practice, the recordings of vital signs have repeatedly been shown to be mostly incomplete.[64, 69-71] The respiratory rate, which is a sensitive predictor of patient deterioration, even is the most commonly unmeasured vital sign on general wards.[72-76] Besides, track-and-trigger tools cannot assist responders with acute decision-making at the bedside when they are a confronted with a deteriorating severely ill patient. While awaiting more experienced help from rapid response teams, these clinicians need to

rely on their pre-existing general knowledge and may fail to notice important details in those complex, ambiguous and stressful moments.[68, 77, 78] Clinical decision support systems can be a valuable solution to help clinicians at the bedside with the collection and interpretation of vital sign measures or to guide their acute care management of deteriorating patients.

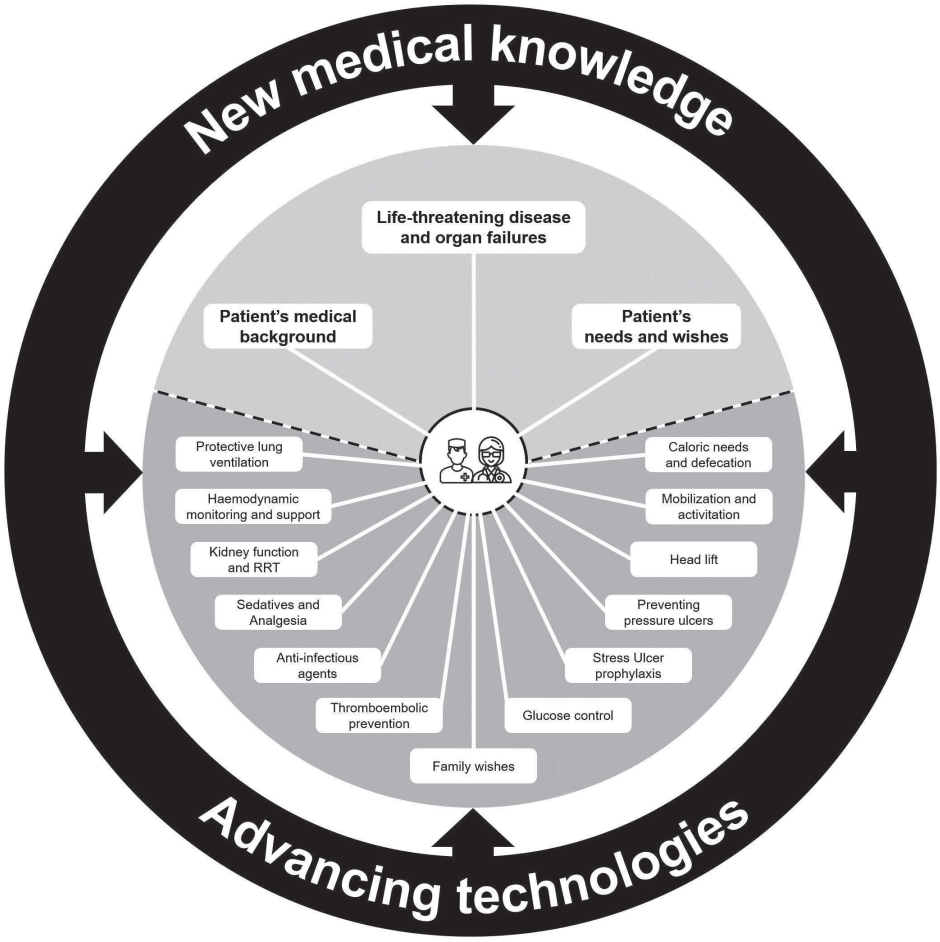


Figure 5. The complexity of the decision-making process in critical care
Healthcare professionals in critical care have to make multiple decisions concerning the management of life-threatening diseases and supportive care processes. New medical knowledge and technologies are constantly altering this decision-making process.

The translation and application of the available medical knowledge and evidence into clinical practice is extremely challenging.[21, 79-81] In the ICU, 20 to 50% of the patients do not receive the recommended care and wide variations exist between institutions. [21, 79] Protocols and checklists have been developed to guide healthcare professionals with the supportive care processes. Several observational studies indicate that better compliance with these protocols and checklists can improve the quality of care.[2, 30, 31, 82] However, in a multicentre randomized controlled trial, a multifaceted quality improvement intervention with daily checklists, goal setting, and clinician prompting did not result in relevant improvements of patient-centred outcomes, even though the adherence to certain care protocols did improve.[35] One of the explanations for this unexpected result might be the difficulty of tailoring the advice of checklists and protocols to the needs of complex critically ill patients with multisystem disease.[83] These patients require a more personalized approach. Therefore, to apply and tailor the best eligible care to each patient's needs, healthcare professionals might appreciate the support of a CDS system with dynamic features, like the personalized digital checklists of TraceBook.

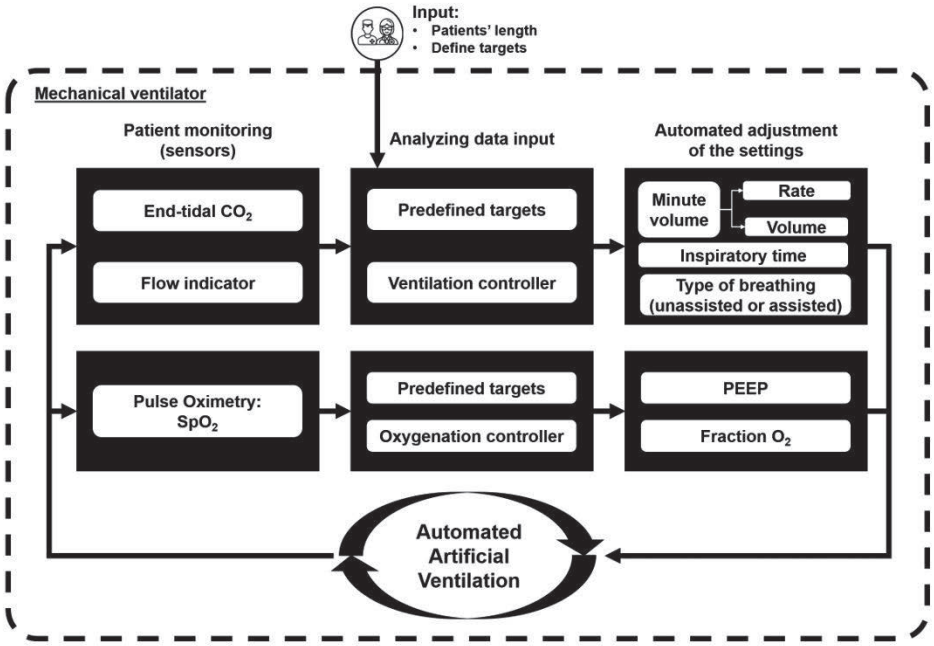


Figure 6. Architecture of a fully closed-loop mechanical ventilation mode
In automated mechanical ventilation various ventilator settings can be adjusted breath-by-breath. The image is a simplified schematic overview of the INTELLiVENT-ASV model.

Another example that illustrates the issues with safety, quality, and knowledge translation in the ICU is the application of lung-protective, low-tidal-volume ventilation in ICU patients. Unsafe ventilator settings affect outcomes of critically ill patients with or without pre-existing lung injury, even in patients who need perioperative mechanical ventilation for a relatively short period, such as cardiac surgery patients.[84-91] Despite this knowledge, studies have repeatedly shown that adherence to lower-tidal-volume ventilation remains difficult, with wide variations in practice persisting while healthcare professionals assumed that they did adhere to this strategy.[21, 81, 92-94] A fully closed-loop ventilation mode (*Figure 6*) might therefore be a solution for improving the adherence to lung-protective ventilation, as it automatically tailors the safest ventilator settings in each breath to the patient's needs.[49, 51, 95-97]

AIM OF THE STUDIES

In this thesis, four different electronic CDS systems will be evaluated to gain insight into the testing and effectiveness of these systems in the area of acute and critical care. The thesis is subdivided into four parts based on the CDS systems' degree of complexity and level of automation (*Figure 1*):

1. **Acute care: A straightforward computerized CDS system without automation (Stage 1)**
   - *Can a Crisis Checklist Application on general wards support the teamwork performance of physicians and nurses and acute care management of a deteriorating severely ill patient?*
2. **Acute care: A CDS system that is based on a simple model which can gather some data automatically (Stage 2)**
   - *Can an automated MEWS system augment the reliable acquisition of vital signs in patients on general wards in order to improve the protocol-adherence and care management of the bedside clinician?*
3. **Critical care: A CDS system based on a complex transparent model with an automated, closed-loop, functionality (Stage 3)**
   - *Can a dynamic clinical checklist of TraceBook improve the adherence of physicians to best eligible practice and their acceptance of checklists during ward rounds in the Intensive Care?*
4. **Critical care: A fully closed-loop automated CDS system (Stage 4)**
   - *Is a fully automated ventilation mode (INTELLiVENT-ASV) able to optimize ventilatory support in patients admitted to Intensive Care?*

REFERENCES

1. Thomassen O, Espeland A, Softeland E, Lossius HM, Heltne JK, Brattebo G. Implementation of checklists in health care; learning from high-reliability organisations. *Scandinavian journal of trauma, resuscitation and emergency medicine* 2011; **19**: 53
2. Vincent JL. Give your patient a fast hug (at least) once a day. *Critical care medicine* 2005; **33**: 1225-9
3. Warshaw G. Introduction: Advances and Challenges in Care of Older People with Chronic Illness *Generations* 2006; **30**: 5-10
4. Benjamin RM. Multiple chronic conditions: a public health challenge. *Public Health Rep* 2010; **125**: 626-7
5. Specialists DAoM. Vision document: Medical Specialist 2025. 2017
6. Medicine Io. *Ensuring Quality Cancer Care Through the Oncology Workforce: Sustaining Care in the 21st Century: Workshop Summary*. Washington, DC: The National Academies Press, 2009
7. Bawden D, Robinson L. The dark side of information: overload, anxiety and other paradoxes and pathologies *Journal of Information Science* 2009; **35**: 180-91
8. Bawden D. Information overload. 2001
9. Osheroff J, Teich, J. , Levick, D. , Saldana, L. , Velasco, F. , Sittig, D. , Rogers, K. , & Jenders, R. Improving Outcomes with Clinical Decision Support: An Implementer's Guide. 2012
10. Middleton B, Sittig DF, Wright A. Clinical Decision Support: a 25 Year Retrospective and a 25 Year Vision. *Yearbook of medical informatics* 2016; **Suppl 1**: S103-16
11. Bezemer T, de Groot MC, Blasse E, et al. A Human(e) Factor in Clinical Decision Support Systems. *Journal of medical Internet research* 2019; **21**: e11732
12. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digit Med* 2020; **3**: 17
13. Greenes RA. Clinical Decision Support. In: Greenes RA, ed. *Clinical Decision Support*. Burlington: Academic Press, 2007; 3-29
14. Musen MA, Shahar Y, Shortliffe EH. Clinical Decision-Support Systems. In: Shortliffe EH, Cimino JJ, eds. *Biomedical Informatics: Computer Applications in Health Care and Biomedicine*. New York, NY: Springer New York, 2006; 698-736
15. Wasylewicz ATM, Scheepers-Hoeks AMJW. Clinical Decision Support Systems. In: Kubben P, Dumontier M, Dekker A, eds. *Fundamentals of Clinical Data Science*. Cham: Springer International Publishing, 2019; 153-69
16. Institute of Medicine (US) Committee on Clinical Practice Guidelines; Field MJ LK, editors. Guidelines for clinical practice: from development to use. 1992
17. Grimshaw J, Russell I. Achieving health gain through clinical guidelines. I: Developing scientifically valid guidelines. *Quality in health care : QHC* 1993; **2**: 243-8
18. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *Bmj* 1999; **318**: 527-30

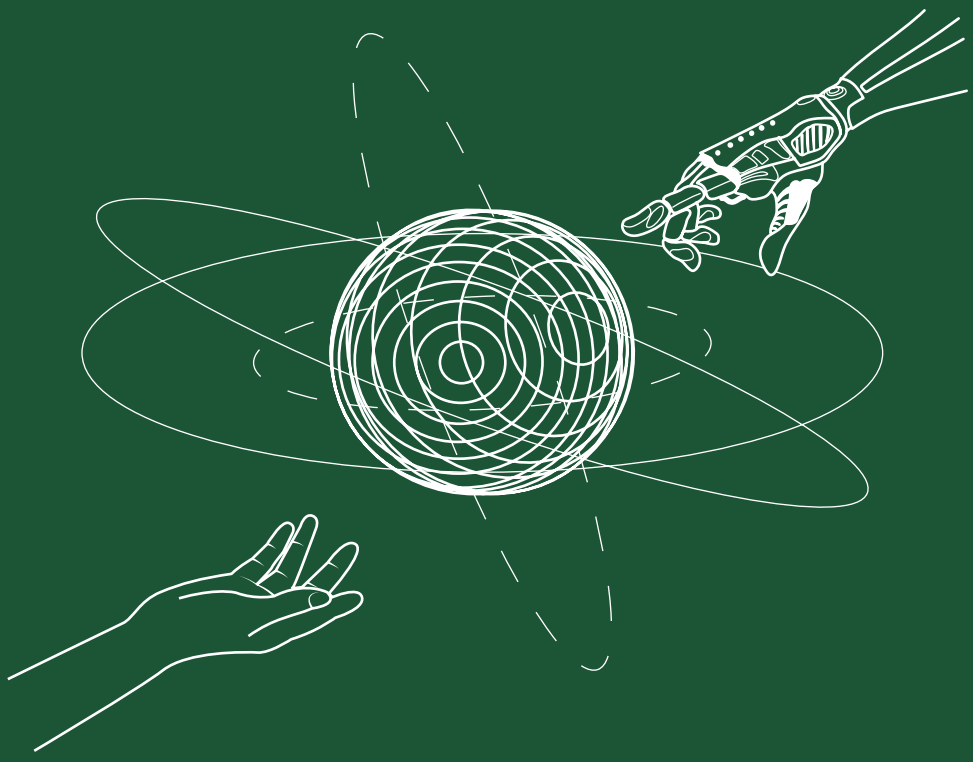
19. Baker M. Clinical guidelines: too much of a good thing. *Annals of the Royal College of Surgeons of England* 2014; **96**: 157-8
20. Upshur RE. Do clinical guidelines still make sense? No. *Annals of family medicine* 2014; **12**: 202-3
21. Ilan R, Fowler RA, Geerts R, Pinto R, Sibbald WJ, Martin CM. Knowledge translation in critical care: factors associated with prescription of commonly recommended best practices for critically ill patients. *Critical care medicine* 2007; **35**: 1696-702
22. Hewson-Conroy KM, Burrell AR, Elliott D, et al. Compliance with processes of care in intensive care units in Australia and New Zealand--a point prevalence study. *Anaesthesia and intensive care* 2011; **39**: 926-35
23. Gawande A. The checklist manifesto : how to get things right. New York: *Metropolitan Books*, 2010
24. Webster CS. Checklists, cognitive aids, and the future of patient safety. *British journal of anaesthesia* 2017; **119**: 178-81
25. Haynes AB, Weiser TG, Berry WR, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *The New England journal of medicine* 2009; **360**: 491-9
26. de Vries EN, Prins HA, Crolla RM, et al. Effect of a comprehensive surgical safety system on patient outcomes. *The New England journal of medicine* 2010; **363**: 1928-37
27. de Vries EN, Prins HA, Bennink MC, et al. Nature and timing of incidents intercepted by the SURPASS checklist in surgical patients. *BMJ quality & safety* 2012; **21**: 503-8
28. Borchard A, Schwappach DL, Barbir A, Bezzola P. A systematic review of the effectiveness, compliance, and critical factors for implementation of safety checklists in surgery. *Annals of surgery* 2012; **256**: 925-33
29. van Klei WA, Hoff RG, van Aarnhem EE, et al. Effects of the introduction of the WHO "Surgical Safety Checklist" on in-hospital mortality: a cohort study. *Annals of surgery* 2012; **255**: 44-9
30. Conroy KM, Elliott D, Burrell AR. Testing the implementation of an electronic process-of-care checklist for use during morning medical rounds in a tertiary intensive care unit: a prospective before-after study. *Annals of intensive care* 2015; **5**: 60
31. Scales DC, Dainty K, Hales B, et al. A multifaceted intervention for quality improvement in a network of intensive care units: a cluster randomized trial. *Jama* 2011; **305**: 363-72
32. Urbach DR, Govindarajan A, Saskin R, Wilton AS, Baxter NN. Introduction of surgical safety checklists in Ontario, Canada. *The New England journal of medicine* 2014; **370**: 1029-38
33. Singer SJ, Molina G, Li Z, et al. Relationship Between Operating Room Teamwork, Contextual Factors, and Safety Checklist Performance. *Journal of the American College of Surgeons* 2016; **223**: 568-80 e2
34. Marshall SD. Helping experts and expert teams perform under duress: an agenda for cognitive aid research. *Anaesthesia* 2017; **72**: 289-95
35. Cavalcanti AB, Bozza FA, Machado FR, et al. Effect of a Quality Improvement Intervention With Daily Round Checklists, Goal Setting, and Clinician Prompting on Mortality of Critically Ill Patients: A Randomized Clinical Trial. *Jama* 2016; **315**: 1480-90

36. Fourcade A, Blache JL, Grenier C, Bourgain JL, Minvielle E. Barriers to staff adoption of a surgical safety checklist. *BMJ quality & safety* 2012; **21**: 191-7
37. Hill MR, Roberts MJ, Alderson ML, Gale TC. Safety culture and the 5 steps to safer surgery: an intervention study. *British journal of anaesthesia* 2015; **114**: 958-62
38. Cresswell K, Majeed A, Bates DW, Sheikh A. Computerised decision support systems for healthcare professionals: an interpretative review. *Informatics in primary care* 2012; **20**: 115-28
39. Wilson A, Duszynski A, Turnbull D, Beilby J. Investigating patients' and general practitioners' views of computerised decision support software for the assessment and management of cardiovascular risk. *Informatics in primary care* 2007; **15**: 33-44
40. Osheroff JA, Teich JM, Middleton B, Steen EB, Wright A, Detmer DE. A roadmap for national action on clinical decision support. *JAMIA* 2007; **14**: 141-5
41. Shortliffe EH, Buchanan BG. A model of inexact reasoning in medicine. *Mathematical Biosciences* 1975; **23**: 351-79
42. Wright A, Sittig DF. A four-phase model of the evolution of clinical decision support architectures. *International journal of medical informatics* 2008; **77**: 641-9
43. Parasuraman R, Sheridan TB, Wickens CD. A model for types and levels of human interaction with automation. *IEEE transactions on systems, man, and cybernetics Part A, Systems and humans : a publication of the IEEE Systems, Man, and Cybernetics Society* 2000; **30**: 286-97
44. Sim I, Gorman P, Greenes RA, et al. Clinical decision support systems for the practice of evidence-based medicine. *Journal of the American Medical Informatics Association : JAMIA* 2001; **8**: 527-34
45. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nature medicine* 2019; **25**: 44-56
46. Kramer HS, Drews FA. Checking the lists: A systematic review of electronic checklist use in health care. *Journal of biomedical informatics* 2017; **71S**: S6-S12
47. Nan S VGP, Korsten H, Vdovják R, Kaymak U. Tracebook: a dynamic checklist support system. *IEEE 27th International Symposium on Computer-Based Medical Systems* 2014: 48-51
48. Nan S, Van Gorp P, Lu X, et al. A meta-model for computer executable dynamic clinical safety checklists. *BMC medical informatics and decision making* 2017; **17**: 170
49. Burns KE, Lellouche F, Nisenbaum R, Lessard MR, Friedrich JO. Automated weaning and SBT systems versus non-automated weaning strategies for weaning time in invasively ventilated critically ill adults. *The Cochrane database of systematic reviews* 2014: CD008638
50. Rose L, Schultz MJ, Cardwell CR, Jouvet P, McAuley DF, Blackwood B. Automated versus non-automated weaning for reducing the duration of mechanical ventilation for critically ill adults and children: a cochrane systematic review and meta-analysis. *Critical care* 2015; **19**: 48
51. Chatburn RL, Mireles-Cabodevila E. Closed-loop control of mechanical ventilation: description and classification of targeting schemes. *Respiratory care* 2011; **56**: 85-102
52. Dellaca RL, Veneroni C, Farre R. Trends in mechanical ventilation: are we ventilating our patients in the best possible way? *Breathe* 2017; **13**: 84-98

53. McConville JF, Kress JP. Weaning patients from the ventilator. *The New England journal of medicine* 2012; **367**: 2233-9
54. Berg M, Langenberg C, vd Berg I, Kwakkernaat J. Considerations for sociotechnical design: experiences with an electronic patient record in a clinical context. *International journal of medical informatics* 1998; **52**: 243-51
55. Berg M. Implementing information systems in health care organizations: myths and challenges. *International journal of medical informatics* 2001; **64**: 143-56
56. Lehoux P. Patients' perspectives on high-tech home care: a qualitative inquiry into the user-friendliness of four technologies. *BMC health services research* 2004; **4**: 28
57. Bates DW, Kuperman GJ, Wang S, et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *Journal of the American Medical Informatics Association : JAMIA* 2003; **10**: 523-30
58. Shortliffe EH, Sepulveda MJ. Clinical Decision Support in the Era of Artificial Intelligence. *Jama* 2018; **320**: 2199-200
59. Lighthall GK, Vazquez-Guillamet C. Understanding Decision Making in Critical Care. *Clinical medicine & research* 2015; **13**: 156-68
60. Needham DM. Patient safety, quality of care, and knowledge translation in the intensive care unit. *Respiratory care* 2010; **55**: 922-8
61. Ghassemi M, Celi LA, Stone DJ. State of the art review: the data revolution in critical care. *Critical care* 2015; **19**: 118
62. Ludikhuize J, Dongelmans DA, Smorenburg SM, Gans-Langelaar M, de Jonge E, de Rooij SE. How nurses and physicians judge their own quality of care for deteriorating patients on medical wards: self-assessment of quality of care is suboptimal*. *Critical care medicine* 2012; **40**: 2982-6
63. Ludikhuize J, Hamming A, de Jonge E, Fikkers BG. Rapid response systems in The Netherlands. *Joint Commission journal on quality and patient safety* 2011; **37**: 138-44, 97
64. Ludikhuize J, Smorenburg SM, de Rooij SE, de Jonge E. Identification of deteriorating patients on general wards; measurement of vital parameters and potential effectiveness of the Modified Early Warning Score. *Journal of critical care* 2012; **27**: 424 e7-13
65. Ludikhuize J, Brunsveld-Reinders AH, Dijkgraaf MG, et al. Outcomes Associated With the Nationwide Introduction of Rapid Response Systems in The Netherlands. *Critical care medicine* 2015; **43**: 2544-51
66. Douw G, Huisman-de Waal G, van Zanten ARH, van der Hoeven JG, Schoonhoven L. Capturing early signs of deterioration: the dutch-early-nurse-worry-indicator-score and its value in the Rapid Response System. *Journal of clinical nursing* 2017; **26**: 2605-13
67. Douw G, Schoonhoven L, Holwerda T, et al. Nurses' worry or concern and early recognition of deteriorating patients on general wards in acute care hospitals: a systematic review. *Critical care* 2015; **19**: 230
68. Massey D, Chaboyer W, Anderson V. What factors influence ward nurses' recognition of and response to patient deterioration? An integrative review of the literature. *Nursing open* 2017; **4**: 6-23

69. Cardona-Morrell M, Prgomet M, Lake R, et al. Vital signs monitoring and nurse-patient interaction: A qualitative observational study of hospital practice. *International journal of nursing studies* 2016; **56**: 9-16
70. Trinkle RM, Flabouris A. Documenting Rapid Response System afferent limb failure and associated patient outcomes. *Resuscitation* 2011; **82**: 810-4
71. Leuvan CH, Mitchell I. Missed opportunities? An observational study of vital sign measurements. *Critical care and resuscitation : journal of the Australasian Academy of Critical Care Medicine* 2008; **10**: 111-15
72. Bellomo R, Ackerman M, Bailey M, et al. A controlled trial of electronic automated advisory vital signs monitoring in general hospital wards. *Critical care medicine* 2012; **40**: 2349-61
73. Cretikos M, Chen J, Hillman K, et al. The objective medical emergency team activation criteria: a case-control study. *Resuscitation* 2007; **73**: 62-72
74. Cretikos MA, Bellomo R, Hillman K, Chen J, Finfer S, Flabouris A. Respiratory rate: the neglected vital sign. *The Medical journal of Australia* 2008; **188**: 657-9
75. Ramachandran SK, Haider N, Saran KA, et al. Life-threatening critical respiratory events: a retrospective study of postoperative patients found unresponsive during analgesic therapy. *Journal of clinical anesthesia* 2011; **23**: 207-13
76. McGain F, Cretikos MA, Jones D, et al. Documentation of clinical review and vital signs after major surgery. *The Medical journal of Australia* 2008; **189**: 380-3
77. Smith GB, Poplett N. Impact of attending a 1-day multi-professional course (ALERT) on the knowledge of acute care in trainee doctors. *Resuscitation* 2004; **61**: 117-22
78. Featherstone P, Smith GB, Linnell M, Easton S, Osgood VM. Impact of a one-day inter-professional course (ALERT) on attitudes and confidence in managing critically ill adult patients. *Resuscitation* 2005; **65**: 329-36
79. Sinuff T, Muscedere J, Adhikari NK, et al. Knowledge translation interventions for critically ill patients: a systematic review*. *Critical care medicine* 2013; **41**: 2627-40
80. Pronovost PJ, Murphy DJ, Needham DM. The science of translating research into practice in intensive care. *American journal of respiratory and critical care medicine* 2010; **182**: 1463-4
81. Schultz MJ, Wolthuis EK, Moeniralam HS, Levi M. Struggle for implementation of new strategies in intensive care medicine: anticoagulation, insulin, and lower tidal volumes. *Journal of critical care* 2005; **20**: 199-204
82. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. *The New England journal of medicine* 2003; **348**: 2635-45
83. Delaney A, Hammond N, Litton E. Checklists and protocols in the ICU: less variability in care or more unnecessary interventions? *Intensive care medicine* 2020
84. Gajic O, Dara SI, Mendez JL, et al. Ventilator-associated lung injury in patients without acute lung injury at the onset of mechanical ventilation. *Critical care medicine* 2004; **32**: 1817-24
85. Brower RG, Matthay MA, Morris A, Schoenfeld D, Thompson BT, Wheeler A. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *The New England journal of medicine* 2000; **342**: 1301-8

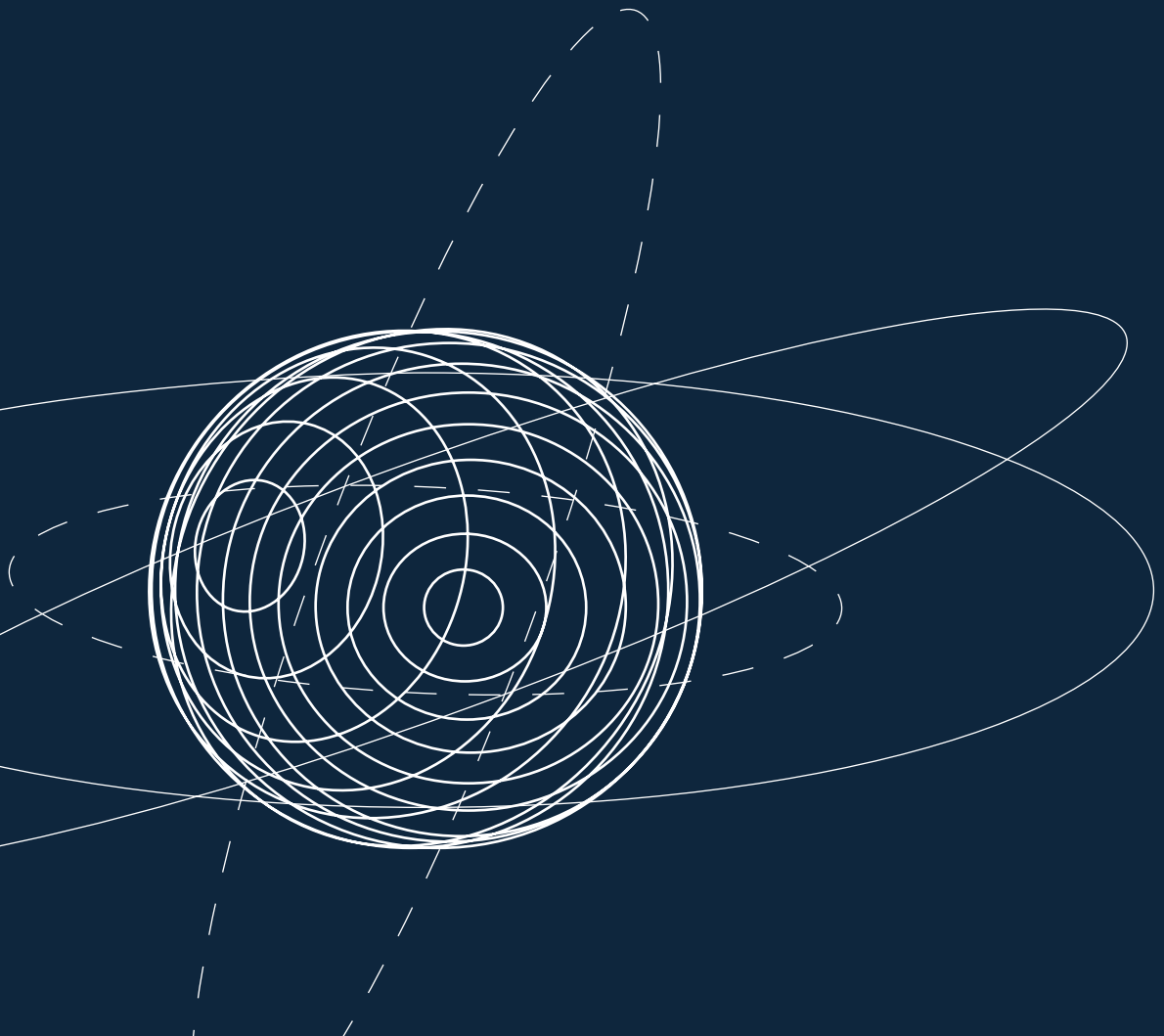
86. Serpa Neto A, Simonis FD, Barbas CS, et al. Association between tidal volume size, duration of ventilation, and sedation needs in patients without acute respiratory distress syndrome: an individual patient data meta-analysis. *Intensive care medicine* 2014; **40**: 950-7
87. Neto AS, Simonis FD, Barbas CS, et al. Lung-Protective Ventilation With Low Tidal Volumes and the Occurrence of Pulmonary Complications in Patients Without Acute Respiratory Distress Syndrome: A Systematic Review and Individual Patient Data Analysis. *Critical care medicine* 2015; **43**: 2155-63
88. Slutsky AS, Ranieri VM. Ventilator-induced lung injury. *The New England journal of medicine* 2013; **369**: 2126-36
89. Putensen C, Theuerkauf N, Zinserling J, Wrigge H, Pelosi P. Meta-analysis: ventilation strategies and outcomes of the acute respiratory distress syndrome and acute lung injury. *Annals of internal medicine* 2009; **151**: 566-76
90. Neto AS, Barbas CSV, Simonis FD, et al. Epidemiological characteristics, practice of ventilation, and clinical outcome in patients at risk of acute respiratory distress syndrome in intensive care units from 16 countries (PRoVENT): an international, multicentre, prospective study. *The Lancet Respiratory medicine* 2016; **4**: 882-93
91. Mathis MR, Duggal NM, Likosky DS, et al. Intraoperative Mechanical Ventilation and Postoperative Pulmonary Complications after Cardiac Surgery. *Anesthesiology* 2019
92. Umoh NJ, Fan E, Mendez-Tellez PA, et al. Patient and intensive care unit organizational factors associated with low tidal volume ventilation in acute lung injury. *Critical care medicine* 2008; **36**: 1463-8
93. Brunkhorst FM, Engel C, Ragaller M, et al. Practice and perception--a nationwide survey of therapy habits in sepsis. *Critical care medicine* 2008; **36**: 2719-25
94. Rubenfeld GD, Cooper C, Carter G, Thompson BT, Hudson LD. Barriers to providing lung-protective ventilation to patients with acute lung injury. *Critical care medicine* 2004; **32**: 1289-93
95. Arnal JM, Wysocki M, Novotni D, et al. Safety and efficacy of a fully closed-loop control ventilation (IntelliVent-ASV) in sedated ICU patients with acute respiratory failure: a prospective randomized crossover study. *Intensive care medicine* 2012; **38**: 781-7
96. Burns KE, Lellouche F, Lessard MR. Automating the weaning process with advanced closed-loop systems. *Intensive care medicine* 2008; **34**: 1757-65
97. Chatburn RL. Classification of ventilator modes: update and proposal for implementation. *Respiratory care* 2007; **52**: 301-23



PART 2

A straightforward computerized CDS system without automation

2.1



Differences in identification of patients' deterioration may hamper the success of clinical escalation protocols

A.J.R. De Bie, C.P. Subbe, R. Bezemer, T. Cooksley, J.G. Kellett, M. Holland, R.A. Bouwman, A.J.G.H. Bindels and H.H.M. Korsten, on behalf of the Crisis Checklist Collaborative

QJM: An International Journal of Medicine 2019

ABSTRACT

Background

Timely and consistent recognition of a "clinical crisis", a life threatening condition that demands immediate intervention, is essential to reduce "failure to rescue" rates in general wards.

Aim

To determine how different clinical caregivers define a "clinical crisis" and how they respond to it.

Design

An international survey.

Methods

Clinicians working on general wards, intensive care units, or emergency departments in the Netherlands, the United Kingdom, and Denmark were asked to review ten scenarios based on common real-life cases. Then they were asked to grade the urgency and severity of the scenario, their degree of concern, their estimate for the risk for death and indicate their preferred action for escalation. The primary outcome was the scenarios with a National Early Warning Score (NEWS) ≥7 considered to be a "clinical crisis". Secondary outcomes included how often a rapid response system (RRS) was activated, and if this was influenced by the participant's professional role or experience. The data from all participants in all three countries was pooled for analysis.

Results

A total of 150 clinicians participated in the survey. The highest percentage of clinicians that considered one of the three scenarios with a NEWS ≥7 as a "clinical crisis" was 52%, while a RRS was activated by <50% of participants. Professional roles and job experience only had a minor influence on the recognition of a "clinical crisis" and how it should be responded to.

Conclusion

This international survey indicates that clinicians differ on what they consider to be a "clinical crisis" and on how it should be managed. Even in cases with a markedly abnormal physiology (i.e. NEWS ≥7) many clinicians do not consider immediate activation of a RRS is required.

INTRODUCTION

Clinical escalation protocols and rapid response systems (RRS) have been developed to reduce the "failure to rescue" rates of hospitalized patient who deteriorate either from their medical condition or a complication of treatment. The goal of these systems is to detect physiological derangements (i.e. the afferent arm) and trigger an effective and appropriate response (i.e. the efferent arm). The efferent arm may consist of the patient's primary carers, or nurse practitioners, or a specialized rapid response team.[1]

Rapid response systems have not consistently reduced in-hospital cardiopulmonary arrests and mortality.[2-4] For these systems to be effective it is essential that there is prompt recognition of deterioration by frontline clinicians, who must also understand how to respond to it appropriately. Therefore, all frontline healthcare practitioners should have an agreed definition of a life threatening condition that demands immediate intervention (i.e. a "clinical crisis") and when care needs to be escalated and additional help called for.[1]

The differences between clinicians in their identification of a "clinical crisis" will undoubtedly hamper the success of any clinical escalation protocols. This study investigated the differences in assessments and recommendations of healthcare practitioners in ten clinical scenarios. The scenarios considered to be a "clinical crisis" was the primary outcome. Secondary outcomes were the clinicians' estimated risk of death within 1 hour, 24 hours and 30 days, and the influence of participants' professional role or experience on their judgements.

METHODS

Study Design

A prospective online hosted survey (SurveyMonkey.com) of ten clinical scenarios was conducted between September 14th and October 28th 2016. Nurses and doctors working on general wards, intensive care units, or emergency units in the Netherlands, United Kingdom, and Denmark were invited by e-mails sent from members of the Crisis Checklist Collaborative. This study was approved by the Medical research Ethics Committees United (MEC-U); reference: W18.193.

Table 1. Short description of the scenarios in the survey

Scenario	Age (years)	Gender	NEWS	Vital signs	Description
1	78	Female	11	BP: 90/40 mmHg HR: 120/min regular RR: 22/min; SpO^2: 94% Temp: 36.5°C; EMV 15	Loss of consciousness: haemorrhagic shock
2	66	Male	3	BP: 150/90 mmHg HR: 103/min regular RR: 15/min; SpO^2: 95% Temp: 38.3°C; E2M4V2 (8)	Sepsis: Cellulitis
3	87	Female	9	BP: 111/55 mmHg HR: 105/min regular RR: 22/min; SpO^2: 91% Temp: 37.9°C; EMV 15	Sepsis: Hospital acquired pneumonia
4	74	Male	3	BP: 95/50 mmHg HR: 110/min irregular RR: 16/min; SpO^2: 97% Temp: 37.1°C; EMV 15	Orthostatic hypotension
5	25	Male	3	BP: 180/75 mmHg HR: 60/min regular RR: 20/min; SpO^2: 99% Temp: 36.5°C; E1M1V1 (3)	Loss of consciousness: subarachnoid bleeding
6	74	Female	0	BP: 130/80 mmHg HR: 82/min regular RR: 16/min; SpO^2: 98% Temp: 37.2°C; EMV 15	Loss of consciousness: hypoglycaemia
7	68	Female	1	BP: 125/80 mmHg HR: 98/min regular RR: 20/min; SpO^2: 98% Temp: 36.8°C; EMV 15	Abdominal pain
8	58	Male	5	BP: 84/50 mmHg HR: 125/min regular RR: 20/min; SpO^2: 98% Temp: 37.8°C; EMV 15	Colon perforation
9	82	Male	5	BP: 110/70 mmHg HR: 138/min irregular RR: 22/min; SpO^2: 95% Temp: 36.6°C; EMV 15	Fast atrial fibrillation
10	72	Male	10	BP: 100/50 mmHg HR: 88/min regular RR: 28/min; SpO^2: 87% Temp: 39.8°C; EMV 15	Sepsis: Community-acquired pneumonia

NEWS, National Early Warning Score; BP, blood pressure; HR, heart rate; RR, respiratory rate; Temp, temperature; EMV, Eye opening, best Motor response, best Verbal response

Selection of scenarios

Members of the Crisis Checklists Collaborative suggest scenarios for the survey based on their own "real life" clinical experience. All submitted scenarios were reviewed by the authors, and a representative sample of these scenarios with a range of physiological abnormalities were included in the survey (*Table 1*; complete survey in Supplemental Digital Content 1). Scenarios were standardised with regards to writing style. Each scenario contained information on the situational context, age, gender, a summary of relevant previous conditions and medication, and a set of vital signs.

The scenarios were stratified according to the degree of the patients' physiological abnormalities using the National Early Warning Score (NEWS) as the template for classification of severity of illness. In the United Kingdom a NEWS of ≥7 mandates escalation to a practitioner with critical care skills.[5]

Survey structure

Each participant was asked to report their professional role, years of experience (<3 years, 3-6 years and >6 years), department of clinical work, and country of residence. Participants were asked to read a scenario followed by four questions:

- Question 1: a classification of the scenario as a "clinical crisis" (i.e. a life threatening condition that demands immediate intervention), "alarming situation", "easy manageable situation" or "non-urgent situation".
- Question 2: the proposed action in the next 30 minutes. The participant could choose to "continue monitoring", "manage treatment on their own", "escalate to a senior colleague" or "initiate the Rapid Response System".
- Question 3: the degree of concern measured against a 10-item Likert scale.

There were no time constraints. Participants were free to stop the questionnaire at any time. Participants could chose to not answer a question, but they were not able to change answers once they had completed the questionnaire.

Outcome

The primary outcome is to report the proportion of clinicians that considered a scenario "clinical crisis". As a secondary outcome the percentage of participants that would activate the RRS within the next 30 minutes will be reported for each scenario. Other secondary outcomes were to compare if participants' professional role or experience affected their judgements, and to report the estimated risk of death within 1 hour, 24 hours and 30 days per scenario.

Statistical analysis

The data from all participants in all three countries was pooled for analysis. Participants were defined as "doctors" if they made clinical decisions and prescribed medication, the remaining participants were all designated as "nurses". Statistical analyses were performed with SPSS version 21 (IBM Corp., Armonk, NY, USA). The distributions of continuous variables were assessed by the Kolmogorov-Smirnov test. The degree of concern, type of immediate action, and estimated risk of death are reported as percentages for each separate scenario. Chi-square tests with Bonferroni corrections were used to test differences between professional role (i.e. doctors or nurses) and years of experience (i.e. ≤6 years or >6 years), as well as the estimated severity of each clinical scenario and the management interventions selected. A Bonferroni corrected alpha level 0.0063 ($p<0.05$; number of analyses = 8) was considered statistically significant.

RESULTS

Participants

The number of respondents ranged from 150 for the 1st scenario to 87 for the 10th scenario; 4,965 of 7,500 possible questions (66.2%) were completed. Half of the doctors and nearly all of the nurses had more than 6 years of experience (*Table 2*).

Table 2. Profiles of the participants (n = 150)

	Netherlands	United Kingdom	Denmark
Participants	70	62	18
Doctors	47.1%	25.8%	94.4%
Nurses	52.9%	74.2%	5.6%
Experience <6 years	35.7%	8.1%	27.8%
Experience ≥6 years	64.3%	91.9%	72.2%

A "clinical crisis".

Three of the ten scenarios had a NEWS ≥7 and two a NEWS <3 (*Figure 1*). None of the participants considered scenario 6 (NEWS = 0) to be a "clinical crisis", yet 79% considered scenario 5, which had a NEWS of 3, to be a "clinical crisis". Although scenario 3 had a NEWS of 9, 13% of participants considered this to be "clinical crisis". Of the two other scenarios with a NEWS ≥7, 40% considered scenario 10 and 52% scenario 1 to be a "clinical crisis ".

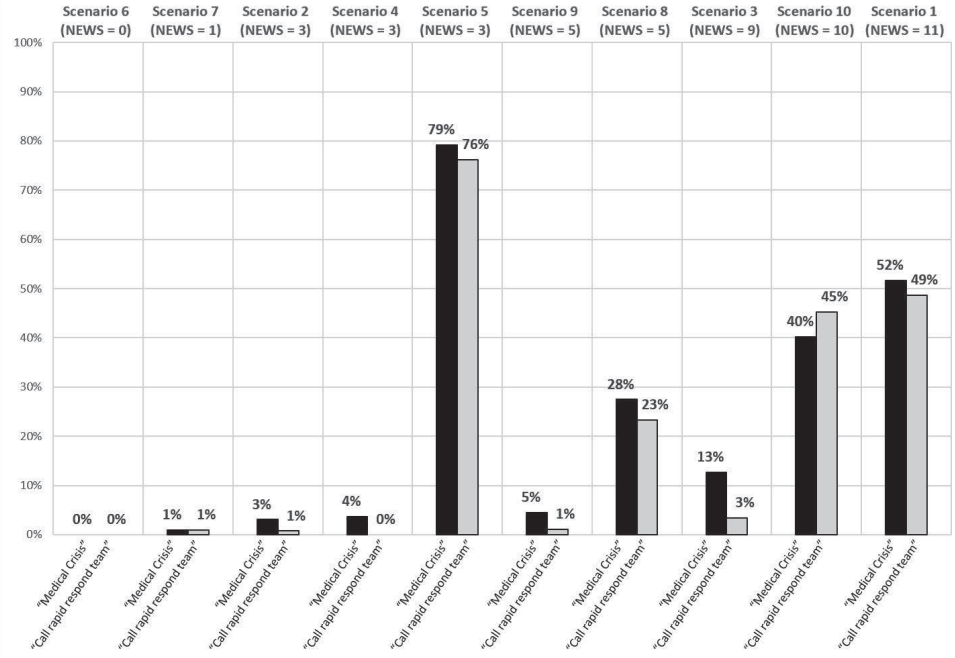


Figure 1. Percentage of participants that classified the scenario as a 'clinical crisis' with the percentage of participants that chose to initiate the rapid response system.
Scenarios are sorted based on increasing NEWS.

Severity of illness and calling the rapid response team within 30 minutes

Both the percentage of participants classifying a scenario a "clinical crisis" and the percentage that chose to activate a RRS within 30 minutes were significantly higher for the scenarios with a NEWS of ≥7 compared to the scenarios with a NEWS of <7 (both p<0.01; *Table 3*). Nevertheless, less than 50% of clinicians would choose to activate a RRS for each of the three scenarios with a NEWS of ≥7 (*Figure 1*). In addition, in scenario 3 (NEWS of 9) 80% of the participants that considered this scenario a "clinical crisis" chose to not activate a RRS (*Table 4*). The estimated risk of death for all time periods were significantly higher in the scenarios with a NEWS of ≥7 compared to the scenarios with a NEWS of <7 (all p <0.01; *Table 3*).

There were some marked discrepancies between the responses of participants. For example, only 13% of participants considered scenario 3 to be a "clinical crisis" and 3% would have activated a RRS, while 40% considered scenario 10 to be a "clinical crisis" and 45% would have activated a RRS, even though both scenarios describe the same underlying disease (i.e. pneumonia) and both had a very high NEWS (*Figure 1*).

Table 3. Comparison of estimated 1 h, 24 h and 30 day mortality, and percentage of participants choosing to activate the RRS within 30-min for scenarios divided into NEWS <7 or 7

	NEWS <7 Median (n; IQR)	NEWS ≥7 Median (n; IQR)	p value (z- or X^2-score)
Estimated mortality within 1 hour	10% (608; 10)	20% (316; 30)	**p<0.001** (z = -7.27)
Estimated mortality within 24 hours	10% (615; 20)	30% (319; 30)	**p<0.001** (z = -11.00)
Estimated mortality within 30 days	20% (636; 30)	50% (328; 40)	**p<0.001** (z = -13.15)
Participants classifying the scenario as a "clinical crisis" (%)	17.0% (n= 122 of 718)	35.9% (n= 127 of 354)	**p<0.001** X^2 = 46.4
Participants choosing to activate the RRS within 30 minutes (%)	14.5% (n= 104 of 718)	32.9% (n= 116 of 353)	**p<0.001** X^2 = 47.8

NEWS, National Early Warning Score; IQR, interquartile range; RRS, rapid response system.

Table 4. Number of participants that classified the scenario as a 'clinical crisis' and decided to not activate the rapid response system (RRS), and of participants that classified the scenario not as a 'clinical crisis' but still decided to activate the RRS. Scenarios are sorted based on increasing NEWS

		Total participants	Classifying scenario as "clinical crisis", but not activating the RRS.		Classifying scenario not as "clinical crisis"; but activating the RRS.	
NEWS	**Scenario**	**n**	**n of n**	**Percentage**	**n of n**	**Percentage**
0	6	97	0 of 0	-	0 of 0	-
1	7	96	0 of 1	0%	0 of 0	-
3	2	131	4 of 4	100%	1 of 127	1%
3	5	106	10 of 84	12%	7 of 22	32%
3	4	109	4 of 4	100%	0 of 105	0%
5	9	88	3 of 4	75%	0 of 84	0%
5	8	91	10 of 25	40%	6 of 66	9%
9	3	118	12 of 15	80%	100 of 103	97%
10	10	87	6 of 35	17%	10 of 52	19%
11	1	149	19 of 77	25%	15 of 72	21%

Scenarios are sorted based on increasing NEWS

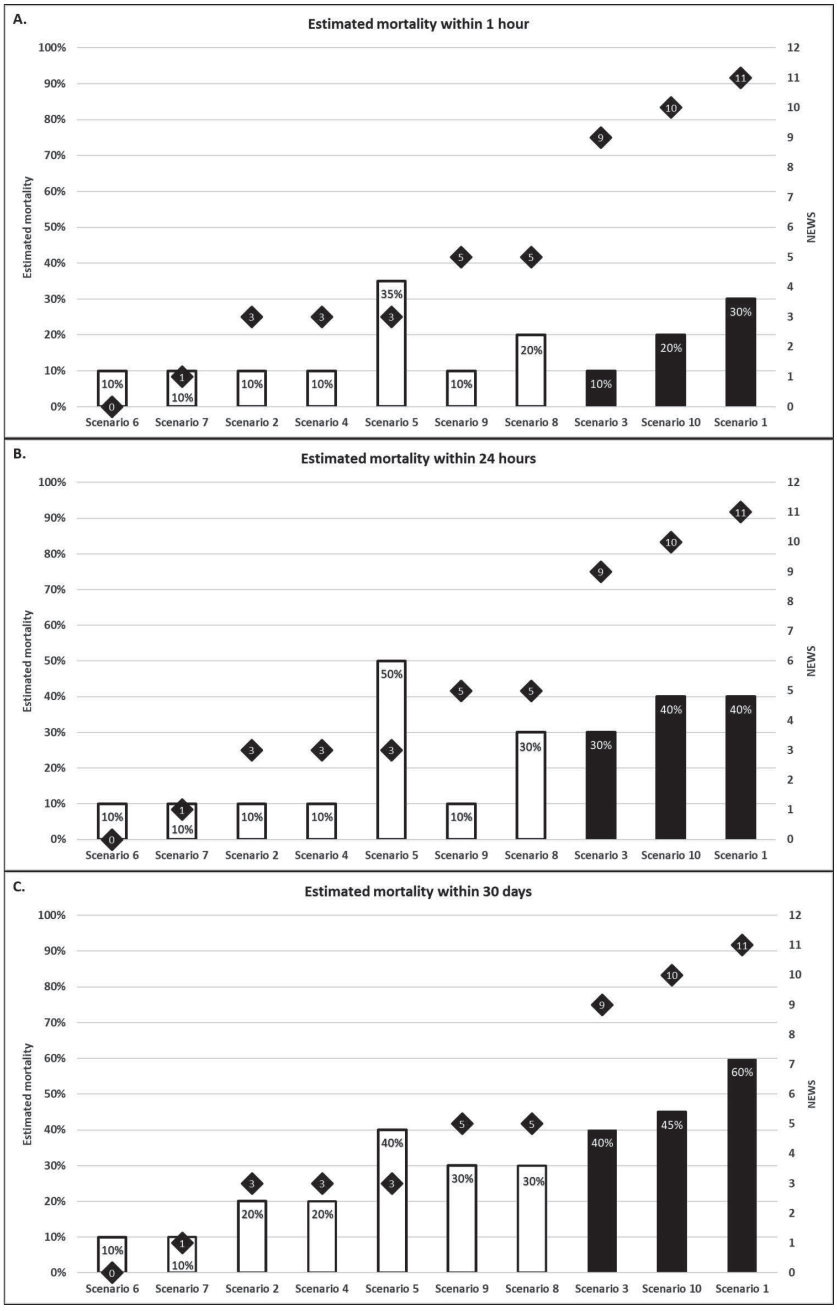


Figure 2. The 1 hour (a), 24 hours (b) and 30 days (c) mortality rate (%) for the patients in each scenario estimated by the responders.

Scenarios in order of increasing NEWS (diamonds); White bars represent scenarios with a NEWS <7 and black bars represent scenarios with a NEWS ≥7.

Estimated risk of mortality

The median estimated risks of dying within 1 hour, 24 hours, and 30 days are reported in *Figure 2*. Only the estimated risk of mortality within 30 days seems to be in line with the increasing NEWS, except for scenario 5 (*Figure 2*). In this scenario of an acute neurological emergency and a NEWS of 3 the median rates were notably higher compared to the scenarios with similar NEWS.

Differences between professions and experience

Considering experience, only scenario two showed that the less experienced participants (experience of ≤6 years) more often classified this scenario as "a situation that could wait" compared to the participants with >6 years of experience ($p<0.05$). In all scenarios experience did not influence the preferred type of action (Supplementary Figure S1-2). Scenario 4 was the only one in which there was a significant difference of opinion between doctors and nurses: more doctors classified the scenario as a situation that could wait. In five scenarios doctors were more inclined to choose a less escalating type of action compared to nurses ($p<0.05$; Supplementary Figure S3-4).

DISCUSSION

A total of 150 healthcare professionals with exposure to RRS, working in hospitals of three European countries assessed ten scenarios with a NEWS ranging from zero to eleven. For scenarios with a NEWS of ≥7 and significant physiological abnormalities 52% was the highest rate of participants that considered one of these scenario as a "clinical crisis". As a result more than 50% did not recommend RRS activation. The scenarios with a NEWS ≥7 or containing neurological abnormalities were more likely to be escalated.

Several studies have examined clinician risk assessments using clinical scenarios to study how acute management decisions are made.[6, 7] This method has been shown to provide valuable insights that would be difficult to obtain in "real life" clinical practice.

Although inappropriate escalation of serious illness has been attributed to a lack of clinical experience[8-14] we have found little or no difference between the assessments of less and more experienced practitioners. Yang et al also found the level of critical care experience had no influence on both paper and physical simulation scenarios, and Thompson et al reported that although critical care experience improved the estimation of the risk, it had no influence on subsequent management.[6, 7] Both these studies were

confined to nurses and, it might be argued, that doctors may be more likely to feel that they have the skills to handle more clinical situations by themselves. However, we found only two scenarios where nurses would escalate care more often than doctors.

Since "Track and trigger" systems and early warning scores have been widely adopted.[2, 3, 15, 16] concern has been expressed that such systems may result in an over-reliance on objective signs and less on intuition and probabilistic reasoning. However, despite having near normal vital signs, the patient with a neuro-surgical catastrophe (Scenario 5) was recognised as a "clinical crisis" requiring a RRS call by the overwhelming majority of participants. Conversely, although the patient with atrial fibrillation and rapid ventricular response (Scenario 9) had abnormal vital signs (NEWS of 5) most participants recognized that the situation was not a "clinical crisis". These two scenarios reflect the capability of practitioners of all grades and experience to recognize deteriorating patient by intuition and interpretation rather than by just numeric vital signs changes.[14, 17, 18]

The threshold for individual clinicians to trigger escalation and their perception of the urgency depends on numerous variables such as physiological derangement, clinical experience, ambient distractions, as well as their confidence, risk tolerance, and sense of personal responsibility.[13, 14, 18] However, these personal biases, opinions and interpretations are difficult to standardize and can lead, especially in stressful situations, to confusion and inconsistency on the shared understanding between caregivers on what constitutes a "clinical crisis" and how to respond to it.

New techniques, such as the "nurse worried score", continuous monitoring, clinical decision support systems that include intelligent, dynamic, patient-specific checklists may help standardize communication and the tracking of how patients respond to interventions.[17-20] This latter aspect is also important to gain more insight into the provided care during a "clinical crisis". Even though these techniques will not replace clinical judgement, they should increase shared understanding between caregivers and, thus, improve the recognition and management of rapidly deteriorating patients.

This study has several strengths and limitations. First, the participants were self-selected clinicians working in acute care, and may not be representative of clinicians working on general wards. Since there were 150 participants it was only possible to subdivide them into two broad professional categories (i.e., doctors and nurses) and two broad experience levels (i.e., <6 and ≥6 years). Although only ten clinical scenarios were tested, they do represent common clinical situations that often prompt calls for help. It is, of course, uncertain that all the participants would act in "real life" in the same way as they

did in the scenarios. Moreover, while the participants' responses may reflect what they would do as solo practitioners, it is possible that the presence of others and/or working in a team might modify their behaviour. Nevertheless, the differences reported here do reflect the variations between caregivers in their interpretation of a patient's state and needs, and the care they feel comfortable providing without calling for more expert help.

CONCLUSION

This international survey shows that the opinions of doctors and nurses differ on illness severity, the need for urgent treatment, and if help is required to deliver it. Even in cases with significant deranged vital signs (i.e NEWS ≥7) many clinicians would not activate a RRS. For successful implementation of escalation protocols a "clinical crisis", and how to respond to it, needs to be more clearly defined.

REFERENCES

1. Ludikhuize J, Hamming A, de Jonge E, Fikkers BG. Rapid response systems in The Netherlands. *Joint Commission journal on quality and patient safety* 2011; **37**: 138-44, 97
2. Ludikhuize J, Brunsveld-Reinders AH, Dijkgraaf MG, et al. Outcomes Associated With the Nationwide Introduction of Rapid Response Systems in The Netherlands. *Critical care medicine* 2015; **43**: 2544-51
3. Maharaj R, Raffaele I, Wendon J. Rapid response systems: a systematic review and meta-analysis. *Critical care* 2015; **19**: 254
4. Alam N, Hobbelink EL, van Tienhoven AJ, van de Ven PM, Jansma EP, Nanayakkara PW. The impact of the use of the Early Warning Score (EWS) on patient outcomes: a systematic review. *Resuscitation* 2014; **85**: 587-94
5. The National Institute for Health and Care Excellence (NICE). Acutely Ill Patients in Hospital. Acutely Ill Patients in Hospital: Recognition of and Response to Acute Illness in Adults in Hospital. London, 2007
6. Thompson C, Bucknall T, Estabrookes CA, et al. Nurses' critical event risk assessments: a judgement analysis. *Journal of clinical nursing* 2009; **18**: 601-12
7. Yang H, Thompson C. The effects of clinical experience on nurses' critical event risk assessment judgements in paper based and high fidelity simulated conditions: a comparative judgement analysis. *International journal of nursing studies* 2011; **48**: 429-37
8. Jackson S, Penprase B, Grobbel C. Factors Influencing Registered Nurses' Decision to Activate an Adult Rapid Response Team in a Community Hospital. *Dimensions of critical care nursing* 2016; **35**: 99-107
9. Gazarian PK, Henneman EA, Chandler GE. Nurse decision making in the prearrest period. *Clinical nursing research* 2010; **19**: 21-37
10. Choudhry NK, Fletcher RH, Soumerai SB. Systematic review: the relationship between clinical experience and quality of health care. *Annals of internal medicine* 2005; **142**: 260-73
11. Bucknall TK. Critical care nurses' decision-making activities in the natural clinical setting. *Journal of clinical nursing* 2000; **9**: 25-35
12. Greaves J, Greaves D, Gallagher H, Steven A, Pearson P. Doctors and nurses have different priorities in using the Modified Early Warning Score protocol. *British journal of anaesthesia* 2016; **116**: 298
13. Roberts KE, Bonafide CP, Paine CW, et al. Barriers to calling for urgent assistance despite a comprehensive pediatric rapid response system. *American journal of critical care* 2014; **23**: 223-9
14. Shearer B, Marshall S, Buist MD, et al. What stops hospital clinical staff from following protocols? An analysis of the incidence and factors behind the failure of bedside clinical staff to activate the rapid response system in a multi-campus Australian metropolitan healthcare service. *BMJ quality & safety* 2012; **21**: 569-75
15. Salvatierra G, Bindler RC, Corbett C, Roll J, Daratha KB. Rapid response team implementation and in-hospital mortality. *Critical care medicine* 2014; **42**: 2001-6

16. Herod R, Frost SA, Parr M, Hillman K, Aneman A. Long term trends in medical emergency team activations and outcomes. *Resuscitation* 2014; **85**: 1083-7
17. Douw G, Schoonhoven L, Holwerda T, et al. Nurses' worry or concern and early recognition of deteriorating patients on general wards in acute care hospitals: a systematic review. *Critical care* 2015; **19**: 230
18. Odell M, Victor C, Oliver D. Nurses' role in detecting deterioration in ward patients: systematic literature review. *Journal of advanced nursing* 2009; **65**: 1992-2006
19. Johnstone CC, Rattray J, Myers L. Physiological risk factors, early warning scoring systems and organizational changes. *Nursing in critical care* 2007; **12**: 219-24
20. De Bie AJR, Nan S, Vermeulen LRE, et al. Intelligent dynamic clinical checklists improved checklist compliance in the intensive care unit. *British journal of anaesthesia* 2017; **119**: 231-8

SUPPLEMENTAL DIGITAL CONTENT 1

Complete survey available at :
https://academic-oup-com.ru.idm.oclc.org/qjmed/article/112/7/497/5369095?searchresult=1#supplementary-data

SUPPLEMENTAL DIGITAL CONTENT 2

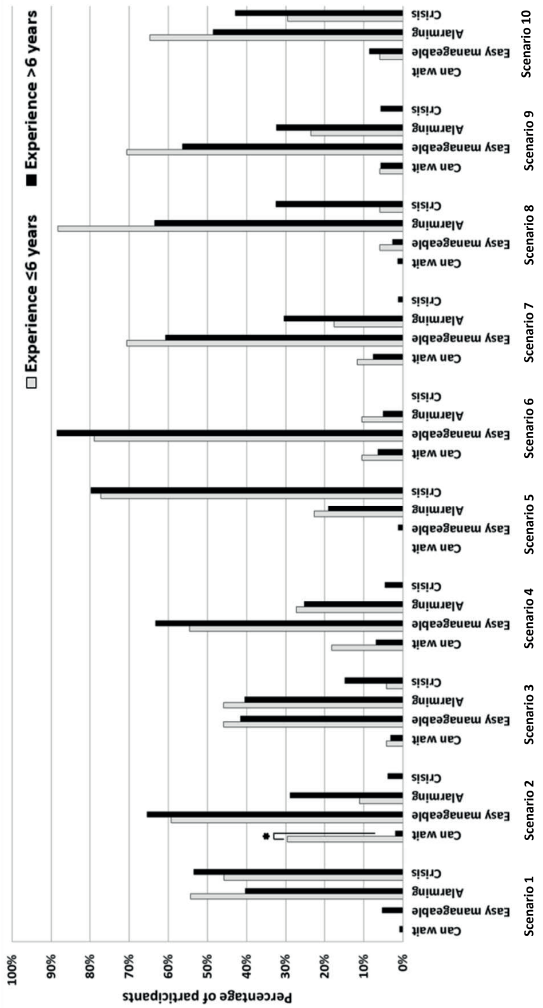
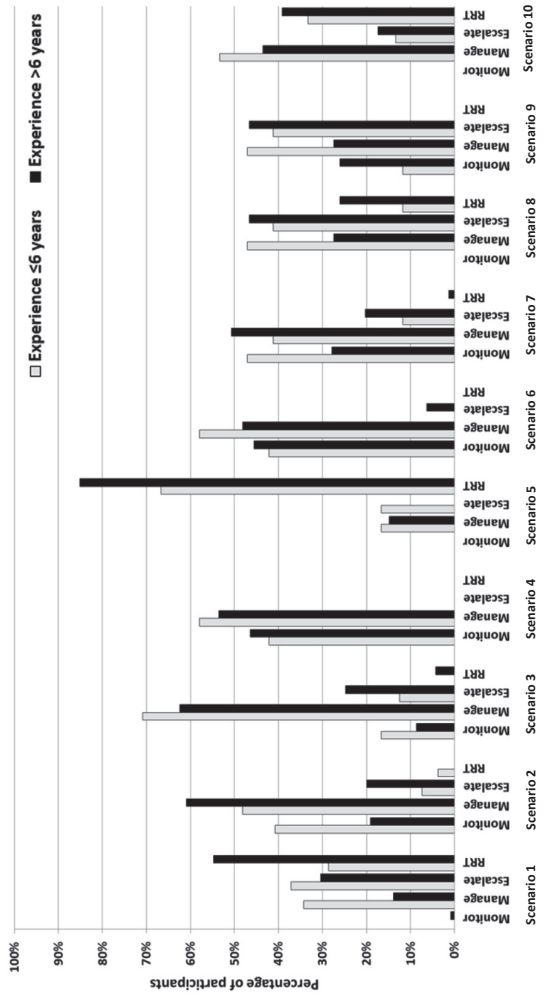


Figure S1. Experience and the classification of the scenarios' severity
*Differences between participants with ≤6 years and >6 years of experience; * p<0.05*



X-axis:
- **Monitor** = *"continue monitoring"*
- **Escalate** = *"escalate to a senior colleague"*
- **Manage** = *"manage treatment on their own"*
- **RRT** = *"initiate the Rapid Response System"*

Figure S2. Experience and the proposed action in the next 30 minutes
*Differences between participants with ≤6 years and >6 years of experience; * $p<0.05$.*

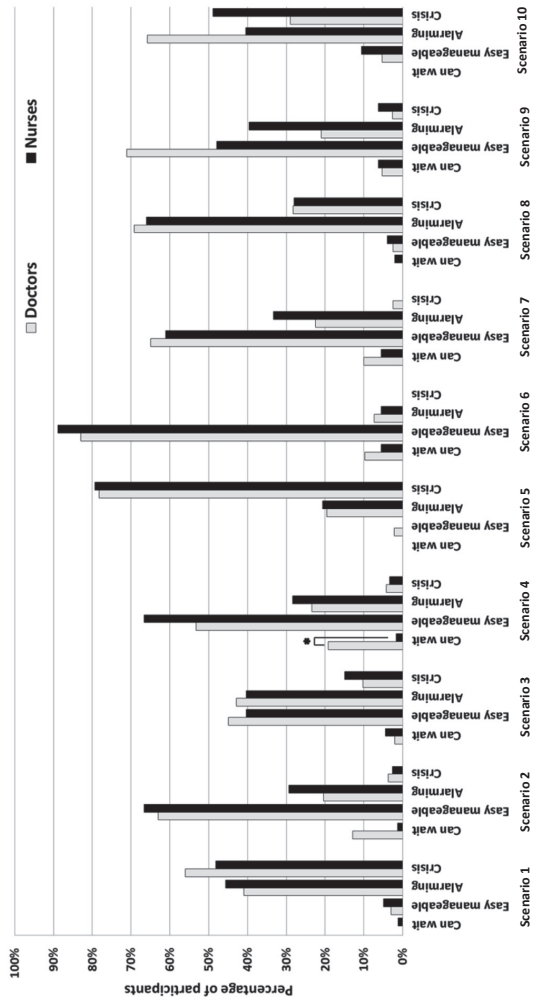


Figure S3. Professional role and classification of the scenarios' severity
*Differences between "doctors" and "nurses"; * p<0.05.*

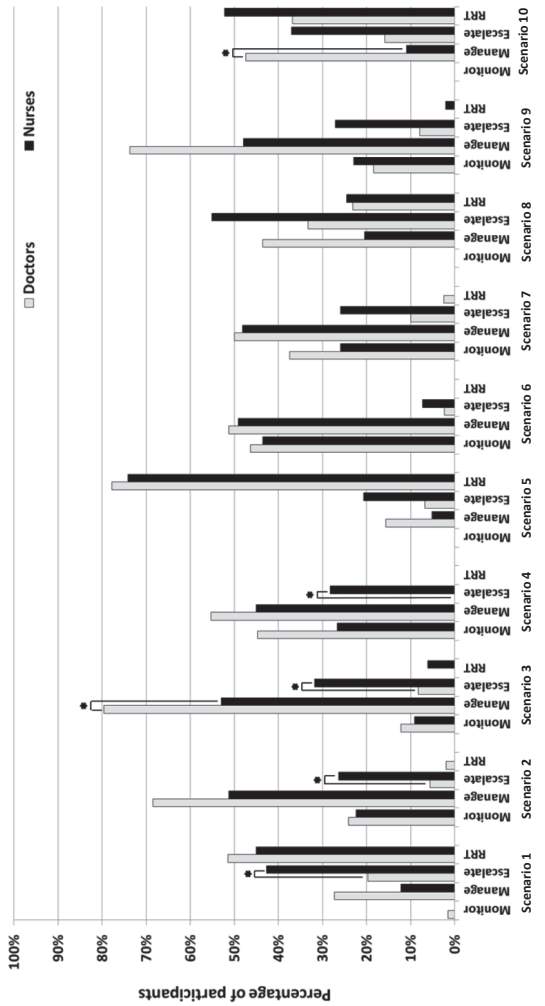
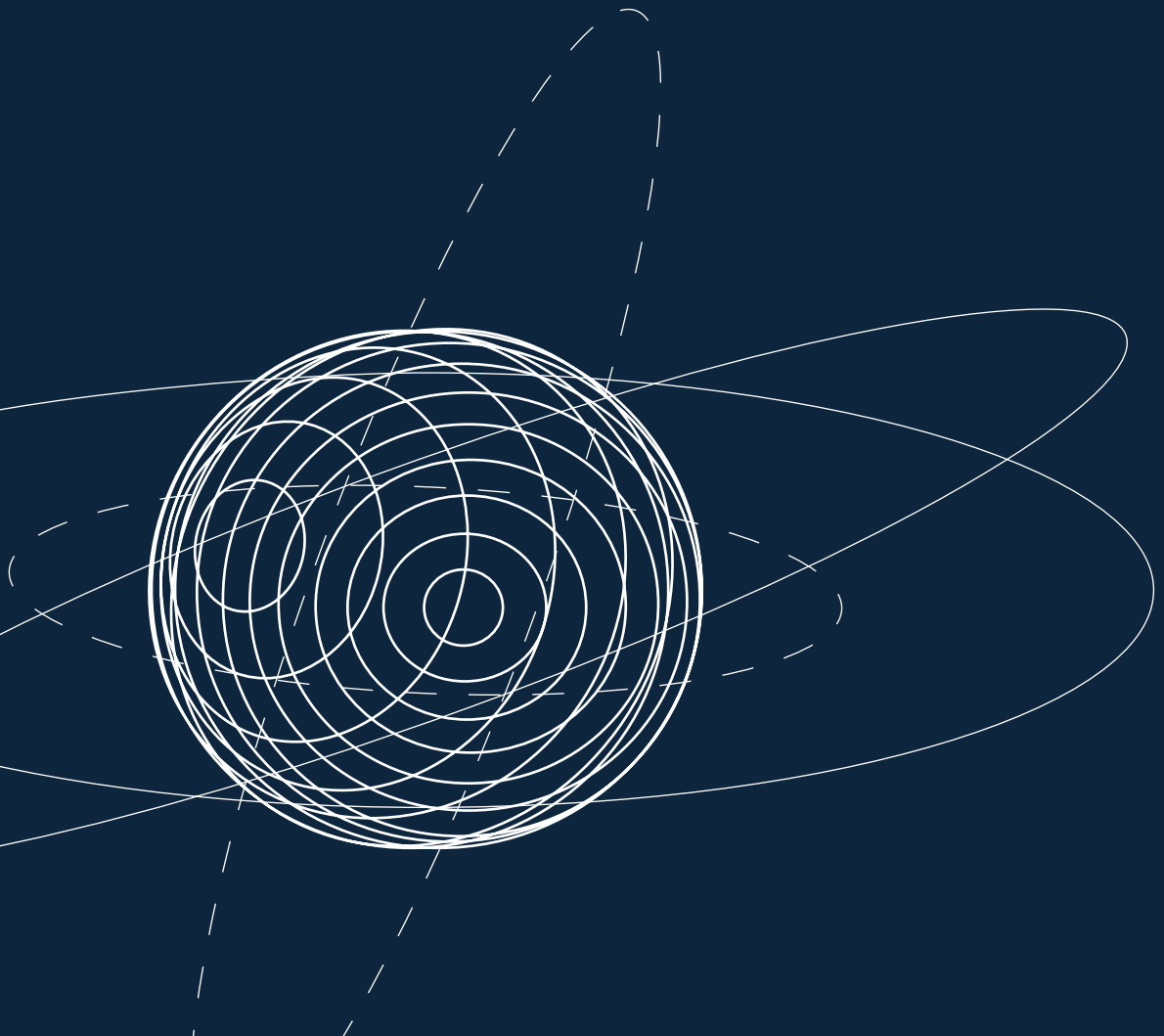


Figure S4. Professional role and the proposed action in the next 30 minutes.
*Differences between "doctors" and "nurses"; * $p<0.05$.*

2.2



Crisis checklists for in-hospital emergencies: expert consensus, simulation testing and recommendations for a template determined by a multi-institutional and multi-disciplinary learning collaborative

C.P. Subbe, J. Kellett, P. Barach, C. Chaloner, H. Cleaver, T. Cooksley, H.H.M. Korsten, E. Croke, E. Davis, A.J.R. De Bie, L. Durham, C. Hancock, J. Hartin, T. Savijn, J. Welch, and on behalf of the Crisis Checklist Collaborative

BMC health services research 2017

ABSTRACT

Background

'Failure to rescue' of hospitalized patients with deteriorating physiology on general wards is caused by a complex array of organisational, technical and cultural failures including a lack of standardized team and individual expected responses and actions. The aim of this study using a learning collaborative method was to develop consensus recommendations on the utility and effectiveness of checklists as training and operational tools to assist in improving the skills of general ward staff on the effective rescue of patients with abnormal physiology.

Methods

A scoping study of the literature was followed by a multi-institutional and multi-disciplinary international learning collaborative. We sought to achieve a consensus on procedures and clinical simulation technology to determine the requirements, develop and test a safe using a checklist template that is rapidly accessible to assist in emergency management of common events for general ward use.

Results

Safety considerations about deteriorating patients were agreed upon and summarized. A consensus was achieved among an international group of experts on currently available checklist formats performing poorly in simulation testing as first responders in general ward clinical crises. The Crisis Checklist Collaborative ratified a consensus template for a general ward checklist that provides a list of issues for first responders to address (i.e. 'Check In'), a list of prompts regarding common omissions (i.e. 'Stop & Think'), and, a list of items required for the safe "handover" of patients that remain on the general ward (i.e. 'Check Out'). Simulation usability assessment of the template demonstrated feasibility for clinical management of deteriorating patients.

Conclusion

Emergency checklists custom-designed for general ward patients have the potential to guide the treatment speed and reliability of responses for emergency management of patients with abnormal physiology while minimizing the risk of adverse events. Interventional trials are needed.

INTRODUCTION

Failure to detect and treat clinical deterioration, either from a medical condition or due to a complication of surgical treatment is a common life threatening problem.[1] Hospitals have introduced Rapid Response Systems (RRS), which use an increasingly standardized evaluation and escalation treatment paradigm to manage patients with physiological derangements.[2,3] In contrast, the efferent limb clinical response is much more variable and ranges from the patient's primary care team, to lone nurse practitioners, to dedicated Rapid Response Teams (RRTs) with intensive care, medical, nursing and allied care providers. The first responders of the efferent limb on general medical and surgical wards will nearly always be an ad hoc assembly of available providers with limited experience in managing common emergency situations.

The publication of 'To Err is Human' in 2000 has prompted a systems-approach towards safe care including applying human factors tools from safety-critical industries, such as aviation and nuclear power, that can be used to mitigate propagation of process failure to systems failures and adverse patient events.[4,5] Checklists have been used effectively as part of routine safety procedures.[6] The introduction of the Safer Surgery Checklist required operating theatre teams in 2009 to change their behaviour in team readiness and has been credited with the reduction of post-operative complications and mortality.[7,8] Studies describing the use of checklists beyond the highly controlled environments of the Intensive Care Unit (ICU) and the operating theatre are rare: the SURPASS trial in the Netherlands demonstrated a slight reduction in mortality associated with the use of multiple checklists during the surgical patient pathway.[9] Implementation of a 'sepsis six' care bundle in general ward areas using a checklist format has also demonstrated only a small reduction in mortality.[10] However, Urbach et al., found that surgical checklists had little impact when clinicians were not involved in checklist design or implementation.[11]

Aviation distinguishes between 'normal', 'non-normal' and 'emergency' checklists.[6] 'Normal' checklists are used as part of standard operating procedures. They include lists used for preparation of a flight or technical checks by maintenance staff. The World Health Organisation's surgical checklist can be seen as a 'normal' checklist. Similar checklists have been used to effectively implement central venous catheter insertion and ventilator associated pneumonia prevention 'care bundles' in many ICUs.[12,13] In these highly controlled settings, checklists seem to have reduced mortality and adverse events and helped to sustain improvements once embedded in clinical practice.[8,12] 'Emergency' checklists deal with uncommon, and unexpected crisis situations likely to have catastrophic outcomes. There are guidelines for the format and content of

'emergency' checklists that specify the recommended colours and typefaces to use.[14] In intensive care and surgery the checklists are intended to be used by several people working together in close partnership.[15]

While Medical Emergency Team call-out criteria and Early Warning Scores have helped to standardize the recognition of deterioration it is not clear how the response could be standardized. We aimed to provide clinicians with rapidly accessible standardized checklists to assist structuring standardized responses to patient deterioration using a checklist format. These explored how checklists could be designed to be used by the patient's 'home' teams and help to structure emergency management and team response to common emergencies during escalation to Rapid Response Teams.

METHODS

Aim

The aim of this study was to develop consensus recommendations on the development and safe testing of a checklist template designed to manage common emergencies that occur on general medical wards by a multi-institutional and multi-disciplinary learning collaborative.

Design: the crisis checklists learning collaborative

A learning collaborative is an innovative and comprehensive approach to multidisciplinary 'action research' that unite researchers, clinicians and policy makers to create a "community of practice".[16] The Crisis Checklists Learning Collaborative came together to create a safe, learning environment for advancing knowledge and promoting best practices related to developing and implementing better care for the deteriorating patient. The group consisted of 32 multidisciplinary experts with over 200 years of combined clinical experience, currently involved in research and clinical practice related to emergency checklists, were invited to participate in a series of consensus meetings. All participants were based at tertiary care medical centres and universities. The invitees attended three face-to-face meetings in Ireland (Dublin April 3-4th, 2014), Wales (Bangor September 5-6th 2014) and England (Manchester January 30-31st 2015).

Our cooperative learning facilitated the accomplishment of a specific end product using the principles of co-design with clinicians and research scientists working together with clinician end-users and patient representatives. Of the 38 participants, 16 were senior medical professionals, nine nursing professionals, six had a technical background while six were in training. Participants included nurse and medical practitioners in the area

of Rapid Response Systems [4], Intensive Care [4], Anaesthesiology [2], acute medical [4] and general ward care [6] as well as a patient representative. The skill set included national and regional program managers for Rapid Response Systems and acute care [4], human factors and patient safety specialists, including those with military and aviation experience, and experts in information technology, quality improvement, systems and graphic design [6] and medical students [2]. Members of the group were from France [1], Germany [3], Ireland [9], Netherlands [2], United Kingdom [17] and the United States of America [2].

Literature search strategy

A scoping study of the literature on checklists and their current use in medical care was performed by three members of the group (JK, CS, PB), prior to the face-to-face meeting, with the aim of summarizing existing research findings and identifying key gaps in the existing literature.[3] We searched for published articles in medical and non-medical literature that assessed the effects of that assessed the effects of checklists. The studies were reviewed for their research design and internal validity. We assessed each study's findings in regard to their effects on patient mortality, morbidity, patient safety, as well as process outcomes. We searched MEDLINE, EMBASE, CRD, for all studies on use of safety checklists. Reference lists of selected articles were searched for potentially relevant studies meeting the inclusion criteria (snowballing). In addition, we used Google search engine using the search words checklist, rapid response team, resuscitation and patient safety. Protocols and publications that outlined safety criteria for use of checklists for deteriorating patients on medical wards were identified and distributed to the group. Additionally, any publication or protocol that a member of the crisis checklist learning collaborative deemed important was circulated prior to the meeting.

Ethics approval

Advice from the Health Research Authority (HRA) was sought with regard to the classification of the study. The HRA classified the collaborative as 'Not Research'. The waiver for informed consent was confirmed by the Bangor Research and Development office.

Workflow learning events

At the first meeting additional knowledge was contributed by participants via presentations from individual group members of any published or unpublished checklist data; further discussions, debate and critique were exchanged in a series of facilitated workshops and focus groups until clear agreement was reached. At the end of the first meeting the following tasks were assigned to designated conference participants:

1. Determine the common clinical situations on general wards for which checklists might be suitable by a further review of the literature
2. Survey experts and active practitioners in rapid response strategies and systems
3. Draft prototype checklists for candidate conditions based on the templates of the Operating Room (OR) Crisis Checklists at www.projectcheck.org/crisis (courtesy to adriadnelab, https://www.ariadnelabs.org)

Following the face-to-face meeting, a summary of the safety criteria for checklists was drafted, and, using an iterative process, was circulated to panel members via email until the group had reached consensus or agreed that they could not reach consensus. Consensus was defined as 100% agreement amongst the group.

The second meeting pilot-tested the checklists for validity and reliability in a high fidelity clinical Simulation Suite at the Ysbyty Gwynedd Hospital, in Bangor, Wales.[16, 18, 19] Test clinical scenarios were undertaken as part of learning events and were administered to teams of volunteer candidates (i.e., doctors, nurses, medical and nursing students) access to medical notes, observation and medication charts and a 'nurse' facilitator delivered the information about their simulated 'patient'. Clinical scenarios were run twice in a randomized fashion, with and without the use of checklists. The performance of teams and individual candidates, with and without, the use checklists was observed, analysed and constructively criticized by the expert participants.

The third collaborative meeting provided the feedback and debate on checklist design, usage, and assessed the role that the clinical culture played in both medical and non-medical settings. Different checklists designs were discussed, piloted, reworked, amended and modified through multiple iterations via discussions, debate and critique in a series of facilitated workshops and focus groups. A consensus on the clinical issues to be addressed by checklists on general medical wards, and the design of the template for these checklists, was ratified by the conference participants. The checklists were edited by a graphic designer and pilot tested with physicians, students and nurses in the Simulation Suite of the Ysbyty Gwynedd Hospital, Bangor, Wales during several sessions in May, and June 2015. Participants self-assessed teamwork, task management, decision making and communication using Likert scales with and without checklists.

RESULTS

The literature search found only two references relating to the use of emergency checklists in operating rooms,[18, 20] and, we found no references related to the use of

checklists and care bundles for emergencies outside intensive care units and operating theatres. We found no published reports evaluating emergency checklist usage on general hospital wards, and no checklists designed for this purpose.

Selection of rapid response team scenarios suitable for checklists

We reviewed published data on the acuity of general ward patients that Rapid Response teams were commonly called on to evaluate. An analysis of 400 calls to a RRT in an Australian Hospital demonstrated that six patient scenario types were responsible for the bulk of call-outs: hypoxia (41%), hypotension (28%), altered conscious state (23%), tachycardia (19%), increased respiratory rate (14%) and oliguria (8%).[19] Clinicians responding to a deteriorating patient could therefore potentially be directed to a limited catalogue of checklists to act upon when treating a deteriorating patient.

A semi-structured survey was designed, piloted, refined, and given to faculty and international specialists in the field from Europe, the US and Australia, to identify candidate conditions for checklists at the International Society for Rapid Response Systems (iSRRS) in Miami in May 2014. A catalogue of candidate conditions amenable to checklists was generated from the survey responses (*Table 1*).

The resulting catalogue of candidate checklists was tested for face validity at the Ysbyty Gwynedd Hospital, a 500 bed facility in the UK. Patients that fulfilled national trigger criteria for a rising National Early Warning Score (NEWS) of 6 or more were reviewed on three general medical wards over a 4 week period.[17] We found 32 patients had new abnormalities, while 68% could be meaningfully allocated to three of the 11 pre-defined scenarios of 'respiratory distress' (38%), 'sepsis' (15%) and 'Altered Loss of Consciousness' (15%).

Table 1. Candidate simulation conditions

Group	Example conditions
Checklists based on Operating Room Crisis Checklists[18]	• Anaphylaxis • Airway • Advanced Life Support scenarios
Interventional crisis	• Gastrointestinal bleed • Myocardial Infarction • Sepsis • Acute Kidney Injury • Fast Atrial Fibrillation
Diagnostic crisis	• Respiratory distress • Un-specifically unwell • Altered mental status
Objective signs of instability	National Early Warning Score (NEWS)[17] level 3, NEWS level 5, NEWS level 7

Simulation testing of emergency checklists templates (Bangor workshop)

The consensus view on currently available checklist formats is that for most providers the use of checklists might bring a 'task-based' rather than a "thought-based" approach to patient management and might result in a failure to seek and consider all available information. For example, in the 'Respiratory Distress' scenario the expected diagnosis of pulmonary embolism was not considered by several candidates. General ward checklists, therefore, need to be designed to prompt comprehensive data gathering and provoke appropriate thought as well as action. Checklist formats similar to the operating theatre checklists,[18] require a team of several responders already at the bedside. However, in a general ward the first-responder is often a lone responder, most likely a registered nurse and/or a junior doctor with limited experience in managing emergencies. The group's consensus view was that emergency checklists for general wards needed to be modified and be consistent with the organizational structure, cultural context and available resources at the time the RRT is called.[21]

Ratification of a checklist template based on consensus opinion (Manchester meeting)

We came to a consensus at the third Collaborative meeting that a general ward checklist should provide a list of key issues for the first responders of the patient's team (home team) to address (labelled the 'Check In'), a list of prompts for further actions or appropriate escalation (labelled 'Stop & Think'), and how they might structure the Rapid Response team intervention, and then list the items required for the safe "hand-off" of patients who had been stabilised and remained on the general ward (labelled 'Check

Out'). We retained the checklist item addressing team leadership ('Who will be the crisis coordinator') from the crisis checklists for the operating room. Our expectation is that this role is either taken up by the most senior clinician or delegated by the same. Twelve candidate checklists were written applying these principles and a graphic designer edited the final version of the checklists for clarity and usability (*Figure 1*).

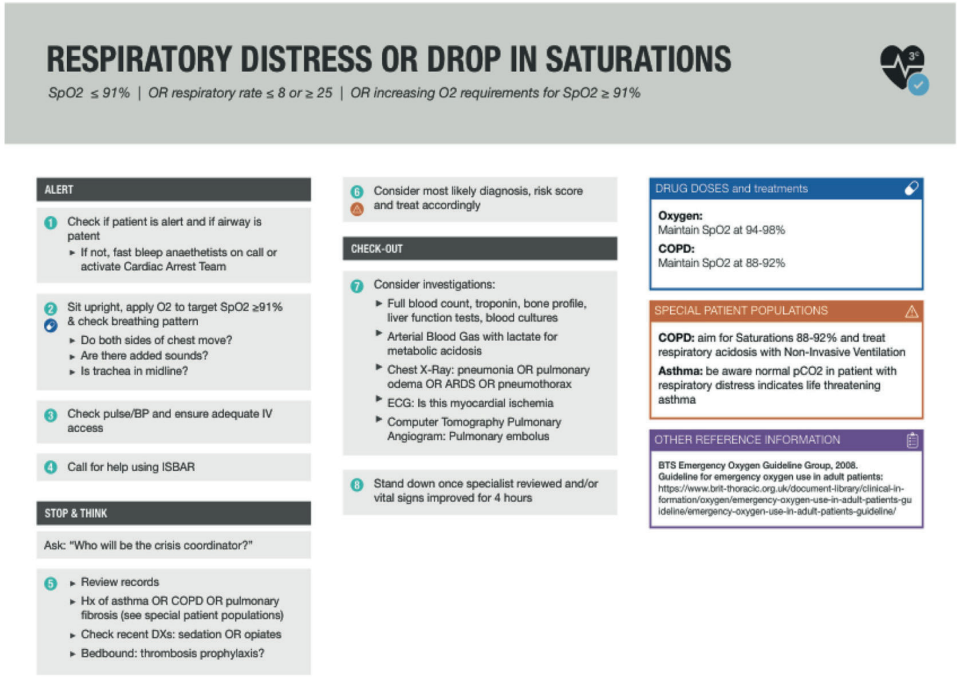


Figure 1. Sample Checklist.
Based on the OR Crisis Checklists at http://www.projectcheck.org/crisis. All reasonable precautions have been taken to verify the information contained in this publication. The responsibility for the interpretation and use of the materials lies with the reader.
Template courtesy to adriadnelab (https://www.ariadnelabs.org/)

Simulation testing of the ratified consensus checklist template

The suite of checklists that addressed the candidate medical emergencies (*Table 1*) was tested by volunteer candidates (i.e. doctors, nurses, medical and nursing students) in simulated environments at five hospitals (Bangor, Cork, Manchester, Rhyl and Eindhoven) in a standardized manner. Several volunteer responders reported usability problems

with the checklists due to lack of familiarity. A short video clip that summarized the rationale and principles of Emergency Checklist was developed to aid training and facility with the simulation mechanics.[22]

When clinical teams were asked to assess their performance during patient management of common simulated emergencies they felt that the use of checklists improved their team work ($p < 0.016$) and communication ($p < 0.01$) and overall performance ($p < 0.034$).

DISCUSSION

The aim of this study was to develop consensus recommendations to provide clinicians about the safe use of emergency checklists to assist in the emergency management of deteriorating patients on general medical wards. Hospitalized patient deterioration continues to challenge healthcare providers with variable outcomes and ongoing preventable harm. Utilizing previous evidence, simulation testing and expert opinion, the learning collaborative group achieved consensus on the best templates to use for RRT teams to assist in structuring patient management when faced with treating deteriorating patients on general medical wards.

This project used an established learning collaborative methodology to gain consensus on developing custom designed and rapidly accessible checklists for ward patients using standard procedures and clinical simulation technology to improve patient management. We found that general wards are qualitatively different from other clinical areas because the first responders must use resources that are available and therefore cannot rely on guidance by specialists. Our experience using a simulated environment suggests that while traditional checklist templates are not appropriate for general ward use, an innovative and flexible template we developed may be of value for the management of the common deteriorating patient by producing rapidly accessible and more reliable responses with improved measures of teamwork.

The systematic assessment of patient physiology at the bedside has led to dramatic reductions in rates of cardiopulmonary arrests.[23-25] Despite this success many instances of abnormal patient physiology do not lead to early activation of a RRT.[26-29] Moreover, even when a RRT team is called key interventions may be missed,[29] possibly as a result of errors in mental modelling and/or an incomplete understanding on how to respond to patterns of abnormal patient physiology.[30] Consequently a significant proportion of patients that trigger a RRT response subsequently generate recurrent "call-outs".[31]

A potential solution for these challenges would be the greater standardisation of RRT activation by routinely using standardized checklists to assist in structuring emergency care management.

While members of the nursing team are usually caring for patients for the duration of their shift most other staff involved may have just transiently entered the ward, and may not have the required competencies. It is incumbent on the first responder to achieve initial stabilisation, best accomplished by using the established airway, breathing and circulation management protocols (i.e. ABCDE).. Therefore, checklists requiring advanced diagnostic and therapeutic skills cannot be activated when a crisis is recognized. More advanced diagnostic and therapeutic interventions can only be provided when more members of the impromptu team arrive. The team's leadership may then need to be re-defined and a reassessment performed using a secondary checklist.

Simulation for testing and training for RRTs and Cardiac Arrest Teams is well established.[32-35] We found testing of checklists in high-fidelity simulation highlighted important differences between patient crises experienced on general wards as compared to templates used elsewhere (i.e. in operating theatres or non-medical settings) due to variable expertise, resources and limited organizational support.

Checklists for emergency management have been used for years by individual clinicians as personal aides de memoire, and health care administrators have encouraged the adoption of checklists in the hope that they will minimize the risk, increase patient safety and cost of litigation.[36] However, as experience with the WHO surgical checklist has demonstrated, the benefits of checklists are only realised when the clinical staff are engaged and they are used to change the dynamics of a team's culture.[37] Medical checklists are more likely to follow a predictable course if they make clinical sense to providers, have clearly defined endpoints and actively engage the teams using them.[38, 39]

Checklists should thus not be regarded as 'magic bullets'. However they can help minimize variation and standardize care, maintain consistency and ensure quality of care resulting in reduced complication rates and lower mortality.[12, 13] Many clinicians, however, worry that checklists may limit their clinical judgment, autonomy, and disrupt professional relationships.[40] These concerns will require significant changes in organisational culture and take time to appreciate and overcome.[41-43] Additionally investment in training will be required to embed the new checklist tools into clinical operations.[44] A vital factor in their successful use is the creation of egalitarian and flattened hierarchical team structures, so that junior team members have 'permission to challenge' and feel psychological safety when raising challenging issues about improving the care processes.[45]

CONCLUSION

The successful implementation of crisis emergency checklists has the potential to improve patient care and outcomes. This study reports on the development of consensus recommendations to provide clinicians with rapidly accessible, standardized emergency crisis checklists to assist in structuring emergency management of patient on the general medical patient wards.

Hospitals are faced with the challenge of improving reliability of their care and patient outcomes especially when treating unstable patients. The concept of emergency crisis checklists is an attractive new addition to the expanding toolkit for continuous quality improvement by clinical teams. RRS crisis emergency checklists are likely to be effective when they are performed as a team routine in the context and readiness for change. An organizational culture that values improving outcomes is essential for sustained uptake and sustained implementation of checklists. The success of checklists will depend on uptake and acceptance by providers, supported by a strongly motivated and committed team ethos. We have drawn on results from a large international learning collaborative team from the US and Europe, comprised of medical and non-medical experts and including specialists from aviation and information technology. Future research required includes systematic evaluation of these recommendations.

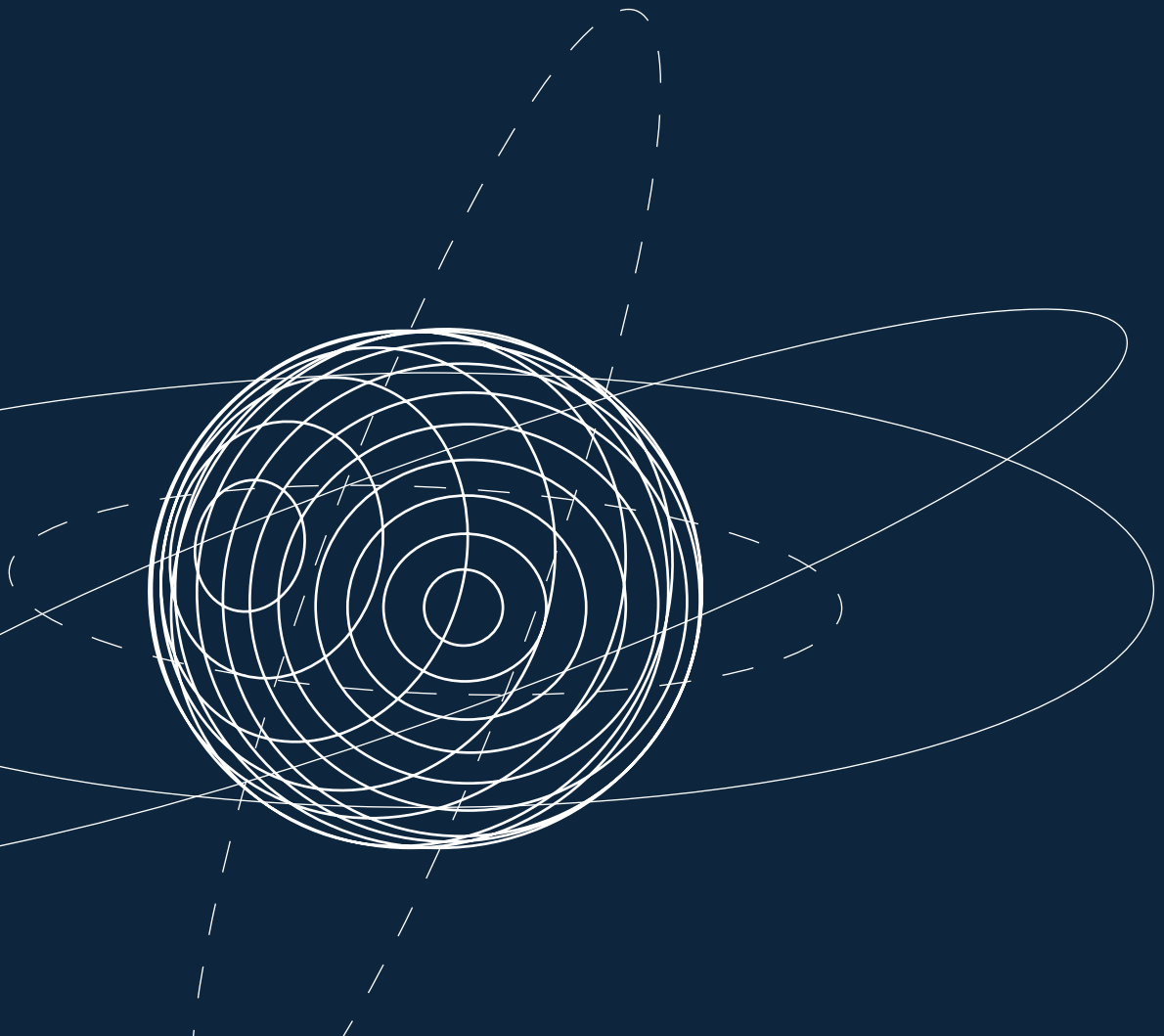
REFERENCES

1. Silber JH, Williams SV, Krakauer H, Schwartz JS. Hospital and patient characteristics associated with death after surgery: a study of adverse occurrence and failure to rescue. *Med Care* 1992; **30**:615–29
2. Lee A, Bishop G, Hillman KM, Daffurn K. The Medical Emergency Team. Anaesth *Intensive Care* 1995; **23**:183–6
3. Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified Early Warning Score in medical admissions. *QJM* 2001; **94**:521–6
4. Kohn LT, Corrigan JM, Donaldson MS, editors. Institute of Medicine (US) Committee on Quality of Health Care in America; To Err is Human: Building a Safer Health System. Washington (DC): *National Academies Press* (US) 2000
5. Amalberti R, Auroy Y, Berwick D, Barach P. Five system barriers to achieving ultra-safe health care. *Ann Inter Med* 2005; **142**:756–64
6. Clay-Williams R, Colligan L. Back to basics: checklists in aviation and healthcare. *BMJ Qual Saf* 2015; **24**:428–31
7. WHO Guidelines for Safe Surgery 2009. Safe surgery saves lives. Geneva: *World Health Organization* 2009; 2009
8. Haynes AB, Weiser TG, Berry WR, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2009; **360**:491–9
9. De Vries EN, Prins HA, Crolla RM, et al. Effect of a comprehensive surgical safety system on patient outcomes. *N Engl J Med* 2010; **363**:1928–37
10. Daniels R, Nutbeam T, Mcnamara G, Galvin C. The sepsis six and the severe sepsis resuscitation bundle: a prospective observational cohort study. *Emerg Med J* 2011; **28**:507–12
11. Urbach DR, Govindarajan A, Saskin R, Wilton AS, Baxter NN. Introduction of surgical safety checklists in Ontario, Canada. *N Engl J Med* 2014; **370**: 1029–38. 11
12. Berenholtz SM, Pronovost PJ, Lipsett PA, et al. Eliminating catheter-related bloodstream infections in the intensive care unit. *Crit Care Med* 2004; **32**:2014–20
13. Resar R, Pronovost P, Haraden C, et al. Using a bundle approach to improve ventilator care processes and reduce ventilatorassociated pneumonia. *Jt Comm J Qual Patient Saf* 2005; **31**:243–8
14. CAP 676: Guidance on the Design, Presentation and Use of Emergency and Abnormal Checklists. *Civil Aviation Authority* (UK) 2006
15. Laughlin PR, Vanderstoep SW, Hollingshead AB. Collective versus individual induction: recognition of truth, rejection of error, and collective information processing. *J Pers Soc Psychol*. 1991; **61**:50–67
16. Wenger E, McDermott RSW. Cultivating Communities of Practice. Boston: *Harvard Business School Press* 2002
17. Jones M. NEWSDIG: The national early warning score development and implementation group. *Clin Med* (Lond) 2012; **12**:501–3

18. Arriaga AF, Bader AM, Wong JM, et al. Simulation-based trial of surgical-crisis checklists. *N Engl J Med* 2013; **368**:246–53
19. Jones D, Duke G, Green J, et al. Medical emergency team syndromes and an approach to their management. *Crit Care* 2006; **10**:R30
20. Ziewacz JE, Arriaga AF, Bader AM, et al. Crisis checklists for the operating room: Development and pilot testing. J Am Coll Surg 2011; **213**:212–7
21. Bosk CL, Dixon-Woods M, Goeschel CA, Pronovost PJ. Reality check for checklists. *Lancet* 2009; **374**:444–5
22. Davis E, Cleaver H. Crisis Checklist Demo https://www.dropbox.com/s/x956nb83bnzbbbg/Crisis%20Checklists.mp4?dl=0 2017
23. Ja G, Ekbom A. Reducing in-hospital cardiac arrests and hospital mortality by introducing a medical emergency team. *Intensive Care Med* 2010; **36**:100–6
24. Jones D, Bellomo R, Bates S, et al. Long-term effect of a medical emergency team on cardiac arrests in a teaching hospital. *Crit Care* 2005; **9**:808–15
25. Merchant RM, Yang L, Becker LB, et al. Incidence of treated cardiac arrest in hospitalized patients in the United States. *Crit Care Med 2011*; **39**:1–6
26. Marshall SD, Kitto S, Shearer W, et al. Why don't hospital staff activate the rapid response system (RRS)? How frequently is it needed and can the process be improved? *Implement Sci* 2011; **6**:39.
27. Yiu CJ, Khan SU, Subbe CP, Tofeec K, Madge RA. Into the night: factors affecting response to abnormal early warning scores out-of-hours and implications for service improvement. *Acute Med* 2014; **13**:56–60.
28. Trinkle RM, Flabouris A. Documenting rapid response system afferent limb failure and associated patient outcomes. *Resuscitation* 2011; **82**:810–4.
29. Hogan H, Healey F, Neale G, Thomson R, Vincent C, Black N. Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study. *BMJ Qual Saf* 2012; **21**:737–45.
30. Durso FT, Sethumadhavan A. Situation awareness: understanding dynamic environments. *Hum Factors* 2008; **50**:442–8
31. Calzavacca P, Licari E, Tee A, et al. Features and outcome of patients receiving multiple Medical Emergency Team reviews. *Resuscitation* 2010; **81**:1509–15
32. Wehbe-Janek H, Lenzmeier CR, Ogden PE, et al. Nurses' Perceptions of Simulation-Based Interprofessional Training Program for Rapid Response and Code Blue Events. *J Nurs Care Qual* 2012; **27**:43–50
33. Devita MA, Schaefer J, Lutz J, Wang H, Dongilli T. Improving medical emergency team (MET) performance using a novel curriculum and a computerized human patient simulator. *Qual Saf Health Care* 2005; **14**:326–31.
34. Parsons SE, Carter EA, Waterhouse LJ, et al. Improving ATLS performance in simulated pediatric trauma resuscitation using a checklist. *Ann Surg* 2014; **259**:807–13

35. Wheeler DS, Geis G, Mack EH, Lemaster T, Patterson MD. High-reliability emergency response teams in the hospital: improving quality and safety using in situ simulation training. *BMJ Qual Saf* 2013; **22**:507–14.
36. Morris JA, Carrillo Y, Jenkins JM, et al. Surgical adverse events, risk management, and malpractice outcome: morbidity and mortality review is not enough. *Ann Surg* 2003; **237**:844–51
37. Fargen K, Mocco J, McConnell D. Surgical checklists: a detailed review of their emergence, development, and relevance to neurosurgical practice. *Surg Neurol Int* 2012; **3**:2
38. Gawande A. The checklist manifesto . 1st ed. *New York: Metropolitan Books* 2009
39. Dixon-Woods M, Leslie M, Tarrant C, Bion J. Explaining Matching Michigan: an ethnographic study of a patient safety program. Implement Sci 2013; **8**:70
40. Catchpole K. The problem with checklists. *BMJ Qual Saf* 2015; **24**:545–9
41. Borchard A, Schwappach DLB, Barbir A, Bezzola P. A Systematic Review of the Effectiveness, Compliance, and Critical Factors for Implementation of Safety Checklists in Surgery. *Ann Surg* 2012; 2**56**:925–33
42. Haynes AB, Weiser TG, Berry WR, et al. Changes in safety attitude and relationship to decreased postoperative morbidity and mortality following implementation of a checklist-based surgical safety intervention. *BMJ Qual Saf* 2011;**20**:102–7
43. Parand A, Burnett S, Benn J, Iskander S, Pinto A, Vincent C. Medical engagement in organisation-wide safety and quality-improvement programmes: experience in the UK Safer Patients Initiative. *Qual Saf Health Care* 2010; **19**:e44
44. Goldhaber-Fiebert SN, Lei V, Nandagopal K, Bereknyei S. Emergency manual implementation: can brief simulation-based or staff trainings increase familiarity and planned clinical use? *Jt Comm J Qual Patient* Saf 2015; **41**:212–20.
45. Edmondson AC. Speaking Up in the operating room: how team leaders promote learning in interdisciplinary action teams. *J Manag Stud* 2003; **40**:1419–52

2.3



Testing the effects of checklists on team behaviour during emergencies on general wards: an observational study using high-fidelity simulation

A.J.R. De Bie, J.J. Dijkmans, N. Todorovac, R. Hibbs, K. Boe Krarup, R.A. Bouwman, P. Barach, M. Fløstrup, T. Cooksley, J. Kellett, A.J.G.H. Bindels, H.H.M. Korsten, M. Brabrand, C.P. Subbe

Resuscitation 2020

ABSTRACT

Introduction

Clinical teams struggle on general wards with acute management of deteriorating patients. We hypothesized that the Crisis Checklist App, a mobile application containing checklists tailored to crisis-management, can improve teamwork and acute care management.

Methods

A before-and-after study was undertaken in high-fidelity simulation centres in the Netherlands, Denmark and United Kingdom. Clinical teams completed three scenarios with a deteriorating patient without checklists followed by three scenarios using the Crisis Checklist App. Teamwork performance as the primary outcome was assessed by the Mayo High Performance Teamwork scale. The secondary outcomes were the time required to complete all predefined safety-critical steps, percentage of omitted safety-critical steps, effects on other non-technical skills, and users' self-assessments. Linear mixed models and a non-parametric survival test were conducted to assess these outcomes.

Results

32 teams completed 188 scenarios. The Mayo High Performance Teamwork scale mean scores improved to 23.4 out of 32 (95% CI: 22.4 – 24.3) with the Crisis Checklist App compared to 21.4 (20.4 – 22.3) with local standard of care. The mean difference was 1.97 (1.34 – 2.6; $p<0.001$). Teams that used the checklists were able to complete all safety-critical steps of a scenario in more simulations (40/95 vs 21/93 scenarios) and these steps were completed faster (stratified log-rank test $\chi^2= 8.0$; $p = 0.005$). The self-assessments of the observers and users showed favourable effects after checklist usage for other non-technical skills including situational awareness, decision making, task management and communication.

Conclusions

Implementation of a novel mobile crisis checklist application among clinical teams was associated in a simulated general ward setting with improved teamwork performance, and a higher and faster completion rate of predetermined safety-critical steps.

INTRODUCTION

Hospital medicine is dealing with increasingly complex patients who often present with multi-morbidity and a combination of conditions that seemingly require conflicting therapeutic strategies.[1-3] At the same time a large proportion of acute and emergency care in hospital medicine is delivered by junior clinicians in the first years of their training.[4, 5] In this setting simple systems using a structured Airway-Breathing-Circulation-Disability-Exposure (ABCDE) approach have become dominant to guide treatment during cardio-pulmonary arrests and peri-arrest situations.[6, 7] They might however be applicable to only a fraction of deteriorating patients and of limited use in solving complex medical problems and interactions between professional groups and disciplines. Very few new approaches to the management of deteriorating patients outside sepsis and cardiac arrests have been developed to support clinicians at the bedside since the first publication of Advanced Life Support courses.[8] The Crisis Checklist Learning Collaborative successfully developed a computerized application with crisis checklists (Crisis Checklist App) using an expert and consensus model to support physicians and nurses in managing acutely deteriorating hospitalized patients on general wards.[9-11]

The objective of this study was to determine the effect of a computerized application with crisis checklists on the acute care management and the teamwork of physicians and nurses who encounter a deteriorating patient. We hypothesized that the Crisis Checklist Application improves the teamwork performance and acute care patient management of clinicians.

METHODS

Study design and setting

We conducted a multicentre simulation study from 1 September, 2017 until 1 December, 2018 in three European hospitals: The Catharina Hospital in Eindhoven in the Netherlands, the Ysbyty Gwynedd in Bangor, Wales, United Kingdom, and Odense University Hospital in Odense, Denmark. This study was undertaken in high fidelity simulation centres representing a typical room on the general ward with interactive mannequins possessing comparable functionalities (i.e., HALL 3201 by Gaumard in Bangor and SimMan 3G by Laerdal in Odense and Eindhoven, respectively) and guided by experienced local simulation facilitators. All facilitators were trained at one of the participating centres.

Participants

Staff members, including final year medical students, nurses and nurse practitioners working on the internal medicine or general surgery services were eligible to participate, as were clinicians working in emergency medicine experienced in general ward clinical coverage.[12] All participating organisations had successfully operationalized rapid response systems. During briefing participants confirmed that they were familiar with their local rapid response system protocol and the ABCDE approach for assessing and managing deteriorating patients.[7, 13]

Design of Crisis Checklists

The crisis checklists were designed by an international learning collaborative group of 32 experts from seven countries in 2016.[11] A computerized application was created for smartphones or tablets to ensure that clinically appropriate prompts are available to the clinician at the bedside on demand. This Crisis Checklist App once opened (available from Apple's App Store[9] and Google's play store[10]) starts by querying five widely accepted ABCDE domains (i.e. Airway, Breathing, Circulation, Disability, Exposure) from which the user can select the domain that best fits the clinical situation.[7, 11] Once the domain is selected, a list of syndromes such as "respiratory distress", "sepsis" or "loss of consciousness" becomes accessible. The syndrome-based checklist provides an easy step-by-step set of instructions including suggestions for proposed diagnostics and potential treatments. Access and links to updated guidelines and resuscitation manuals are also provided (*Figure 1*).

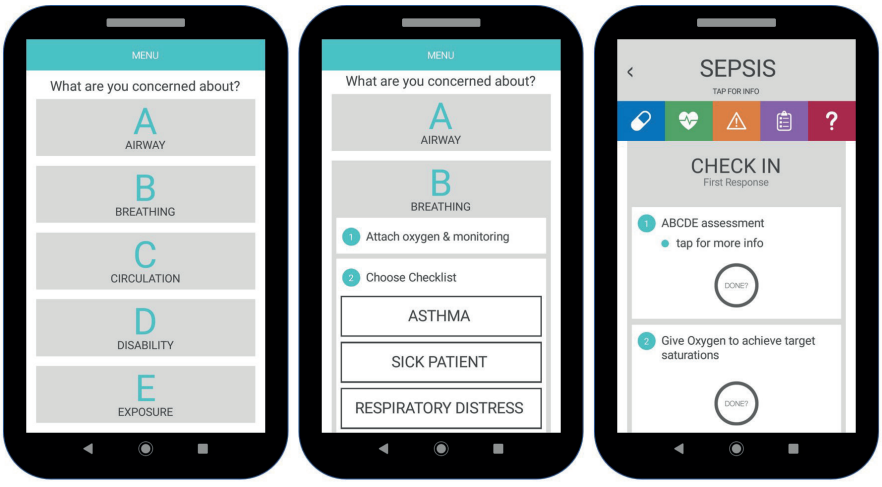


Figure 1. Screenshots of the Crisis Checklist App

Development of scenarios

Two groups of three scenarios each (scenarios 1–3 and 4–6) were created based on previously validated training scenarios from a study evaluating the perceptions of risk about medical emergencies on general medical wards.[14] The patients in the scenarios had comparable clinical presentations: "respiratory distress", "sepsis" and "loss of consciousness". A scenario in group 1 was matched to a scenario of group 2 based on the syndrome-based checklists and on comparable measures of severity of illness (APACHE IV score; *Table 1*). The case vignettes and corresponding scripts are described in detail in Appendix 1.

Simulation protocols and procedures

Participants were divided into teams based on their availability to participate and each team consisted of one "junior nurse" (<3 years clinical experience), one "senior nurse" (>3 years clinical experience), and at least one "medical practitioner" (a resident, final year medical student or nurse practitioner). The teams were randomly assigned (www.random.org) to perform all the scenarios according to two different schedules and in a different order to ensure that the measured effect was not simply a result of the content of the scenarios. Schedule A starts with scenarios 1–3 followed by matched scenarios 4–6; Schedule B starts with scenarios 4–6 first and then flips to scenarios 1–3. Three scenarios were performed initially without the checklists (i.e. the local standard of care). The team received a brief tutorial on how to use the Crisis Checklist App and three matched scenarios were subsequently performed.

Each clinical scenario begins with the junior nurse receiving a briefing of a scripted medical history and patient admission details. The junior nurse is asked to evaluate and treat the patient. The junior nurse can then elect to involve the senior nurse by using a paging device. The nurse is entitled to contact the medical resident at any time. The scenarios were stopped when the team indicated they had completed the scenario, or alternatively, after 15 minutes.

Participants were asked to self-assess their performance with a structured questionnaire after each scenario. After the completion of all scenarios, the voluntary participants had the opportunity to review the scenarios in a facilitator-guided debriefing directed by the experienced local simulation facilitators. These debriefings generally lasted between 30 - 60 minutes without a specific script in which the participants could diffuse, explore and discuss their feelings and experienced events with video reviews to facilitate the discussion.

Table 1. Summary of the study scenarios

	Group 1			Group 2		
Scenario	**1**	**2**	**3**	**4**	**5**	**6**
Sex	Male	Female	Male	Male	Female	Female
Age (years)	55	78	25	66	63	78
Pathology	Massive pulmonary embolism	Community acquired pneumonia	Subarachnoid haemorrhage	Aspiration	Bowel perforation	Hypovolemic shock
Syndrome-based checklist	Respiratory distress	Sepsis	Loss of consciousness	Respiratory distress	Sepsis	Loss of consciousness
NEWS	11	11	5	12	5	11
APACHE IV scores at presentation	82	93	115	81	82	120
Number of defined safety-critical steps	8	9	11	10	9	11
Times a scenario was completed	32	32	32	32	31	30
The number of analysed scenarios	32	32	32	32	31	28
Mean duration of scenario (sec)	777	791	562	714	732	728

Data collection

The data was obtained from each simulation using standardized data sheets. The data collectors received uniform training over two days and supervision from the primary investigators (ADB, CS, and MB) in the identification and classification of complications and process measures. All simulation scenarios were recorded on multi-screen synchronized videos which were independently reviewed by six pairs of ten observers (ADB and JD, AT and NT, MB and NT, NJ and NT, CH and JW, NL and SD). Per team, each video recorded scenario was assessed by the same pair of observers and their scores were averaged per scenario.

Study outcomes

Primary outcome

The primary outcome of the study was team performance as measured by observers using the Mayo High Performance Teamwork scale. The Mayo Scale is a tool to assess teamwork performance in simulated clinical emergencies and has been used in contexts representing critical anaesthesia management or emergency response team situations, critical care events on a Paediatric Intensive Care Unit, and in trauma resuscitation.[15-17] The Mayo scale has been validated in simulations that were designed to replicate emergency response team situations with a deteriorating patient, which is comparable to the ward emergencies in this study.[15] The tool contains 16 items that address four categories of behaviours: communication, leadership, situation awareness, and decision making. Each of these items can be rated from 0 (never/rarely) to 2 (consistently) or not applicable, resulting in a maximum score of 32 points.

Secondary outcomes

The secondary outcomes were the quality of clinical decision making by the team using a 5-point Likert scale in terms of whether it hindered or enhanced teamwork. A self-assessment questionnaire completed by each participant after a scenario and the FoNTS matrix was assessed by the observers (Appendix 1).

The FoNTS matrix tool was designed to assess critical non-technical skills of inexperienced doctors being trained to manage simulated deteriorating patients. The scale covers four domains: Situational Awareness, Decision Making, Task Management, and Teamwork. The domains are rated according to specified criteria from 1 (poor) to 4 (good), with a maximum total score of 16 points (Appendix 1).[18]

Other secondary outcomes were the proportion of omitted pre-defined safety-critical steps, and the time from the start of the scenario to performance of safety-critical steps.

The number of defined safety-critical steps per scenario varied from 8 to 11 (*Table 1* & Appendix 1). The Safety-critical steps were defined by published guidelines, e.g. the NICE guidelines, and expert opinion in a series of consensus meetings, and were tested in a feasibility study.[11]

Ethics approval

The local Research Ethics Committees approved the study in each participating hospital. BASIC was registered in the Netherlands Trial Registry (NTR6666). Participants signed an informed consent agreeing to be videotaped.

Statistical analysis

Power-calculations based on the previous pilot study indicated the need for a sample size of 170 scenarios (Appendix 1).[11] The statistical analyses were performed using IBM SPSS (Statistical Package for Social Sciences, version 22.0; IBM Corp, Armonk, NY, USA) and Genstat® (Windows 13th Edition; VSN International, Hemel Hempstead, UK). The inter-observer reliabilities between the six pairs of observers were calculated among each video recorded scenario. Inter-observer reliability for the Mayo Scale and FoNTS Matrix are described using a percentage of agreement and the linear weighted Cohen Kappa coefficient. The weighted Cohen Kappa coefficients were interpreted by the Landis and Koch scale.[19, 20]

Linear mixed effects modelling (LMM) fit by restricted maximum likelihood was conducted to account for the non-independent repeated measures within the participating teams. The LMMs were applied to assess the significance of the intervention in explaining variations in the MAYO scale, the FoNTS matrix and the self-assessment questionnaire scores, and the percentages of omitted safety-critical steps. In these models the following factors were included: checklist usage (fixed; yes or no), type of scenario (fixed; scenario 1 to 6), the randomized group (fixed; schedule A and B), and each team (Random; team 1 to 32). The distribution of residuals was assessed with the analyses of histograms and P-P plots. Since regression-based models can be sensitive to variables that are correlated, the variance inflation factors (VIF) for all factors used in the model were calculated to check for multi-collinearity (Appendix 1 in Supplementary data).

The likelihood that teams completed all the safety-critical steps faster with the use of the checklists was modelled using the non-parametric survival analysis from Genstat® since this analysis allows both interval and right censoring of the event of interest over time and stratification. The event was taken to be the completion of all the safety-critical

steps and the time component was deemed to be the amount of seconds from the start of the scenario until the completion of the final step. In instances where all steps were not completed or timepoints were unavailable, then observations were treated as being right censored and time was equated to a 1000 seconds (the longest time for a scenario). Hypothesis tests between the non-parametric survival curves were based on log-ranks tests that were stratified with the scenarios as stratum since the numbers of safety-critical steps in each scenario differ.

We conducted an additional post-hoc subgroup analysis with a t-test to compare the effect of checklist usage on percentages of omitted safety-critical steps for teams that were led by a senior student versus those that were led by a qualified clinician (doctor or advanced nurse practitioner).

All reported p-values are two-sided with a $p<0.05$ considered significant, and no adjustments were made for multiple comparisons. 95% confidence intervals were constructed. The guidelines for Strengthening the Reporting of OBservational studies in Epidemiology (STROBE) with the Simulation-Based Research Extensions were used in preparing this article.[21, 22]

RESULTS

Participants and scenarios

We enrolled 101 volunteers for the study. The participants were assigned to 32 teams: 18 teams were randomized to schedule A and 14 teams to schedule B. There were no significant differences in the participants' characteristics between the schedules (*Table 2*). Two scenarios could not be analysed due to video recording malfunctions and one team was unable to complete two scenarios due to an unexpected clinical duty that intervened. Of the 188 reviewed scenarios, 93 scenarios were completed without checklists and 95 scenarios were completed with checklists.

Mayo scale

The mean Mayo scale score for the observers was higher (i.e. better teamwork) for scenarios that were performed with the checklists as compared to scenarios without them (22.4 (95% CI =22.4 – 24.3) vs. 21.4 (95% CI =20.4 – 22.3) out of 32 points with a mean difference of 1.97 (95% CI =1.34 – 2.60); $p<0.001$; *Table 3*).

Table 2. Characteristics of the Study Participants in Schedule A and B.

	Schedule A (n=57)	Schedule B (n=44)
Gender		
Female	38 (67%)	26 (82%)
Male	19 (33%)	8 (18%)
Occupation		
Resident	13 (23%)	7 (16%)
Nurse practitioner	2 (4%)	1 (2%)
Junior Nurse	19 (33%)	12 (27%)
Senior Nurse	13 (23%)	9 (21%)
Senior Student	10 (17%)	15 (34%)
Highest degree of participating staff members		
Community college	8 (14%)	6 (14%)
Professional college	18 (32%)	12 (27%)
University	24 (42%)	26 (59%)
Experience		
≤ 3 years	30 (53%)	27 (61%)
> 3 years	27 (47%)	17 (39%)
Experience with SaferSurgery Checklist		
Yes	7 (12%)	7 (16%)
No	37 (65%)	24 (55%)
Training in acute health care		
ABCDE approach	57 (100%)	44 (100%)
Immediate Life Support	38 (67%)	31 (71%)
Advanced Cardiovascular Life Support	6 (11%)	3 (7%)
Advanced trauma life support	3 (5%)	3 (7%)
Acute Life Threatening Events Recognition and Treatment	4 (7%)	1 (2%)
Area of work		
Internal medicine	11 (19%)	10 (23%)
Intensive care	9 (16%)	11 (25%)
Cardiology	1 (2%)	1 (2%)
Surgery	10 (17%)	-
Senior student	21 (37%)	21 (48%)
Other	5 (9%)	1 (2%)

Table 3. Linear mixed model results of observers' MAYO scale and FoNTS matrix, and the participants' self-assessment.

	Scenarios without checklists Mean (95% CI)*	Scenarios with checklists Mean (95% CI)*	Mean Difference	Standard Error	Paired t-test	p-value	95% Confidence Interval	
							Lower Bound	Upper Bound
Mayo high performance resource scale								
Overall score (out of 32)	21.37 (20.41 – 22.33)	23.33 (22.37 – 24.29)	1.97	0.32	6.18	<0.001	1.34	2.60
Subgroup analysis: person in lead Overall score (out of 32)								
Resident (n=19 in 111 scenarios)	21.83 (20.48 – 23.19)	24.22 (22.86 – 25.57)	2.39	0.38	6.29	<0.001	1.63	3.14
Nurse practitioners (n=3 in 17 scenarios)	20.98 (18.38 – 23.58)	25.54 (23.42 – 27.67)	4.56	1.49	1.29	0.01	1.20	7.92
Senior student (n=10 in 60 scenarios)	20.19 (18.98 – 21.41)	20.98 (19.76 – 22.20)	0.78	0.50	1.59	0.12	-0.21	1.78
FoNTS matrix								
Domains								
Situational awareness (out of 4)	2.83 (2.66 – 2.99)	3.15 (2.98 – 3.32)	0.32	0.07	4.92	<0.001**	0.19	0.45
Decision making (out of 4)	2.81 (2.63 – 2.98)	3.14 (2.96 – 3.32)	0.33	0.07	4.83	<0.001**	0.20	0.47
Task management (out of 4)	2.96 (2.77 – 3.14)	3.17 (2.99 – 3.36)	0.22	0.06	3.42	0.001**	0.09	0.34
Team working (out of 4)	3.07 (2.93 – 3.22)	3.44 (3.29 – 3.59)	0.37	0.07	5.51	<0.001	0.24	0.50
Overall score (out of 16)	11.66 (11.08 – 12.25)	12.89 (12.30 – 13.48)	1.23	0.18	6.75	<0.001**	0.87	1.59
Subgroup analysis FoNTS matrix: person in lead; Overall score (out of 16)								
Resident (n=19 in 111 scenarios)	12.06 (11.34 – 12.78)	13.22 (12.50 – 13.94)	1.16	0.24	4.91	<0.001**	0.68	1.63
Nurse practitioners (n=3 in 17 scenarios)	11.79 (10.64 – 12.95)	13.13 (12.18 – 14.07)	1.33	0.66	2.02	0.07	-0.16	2.80
Senior student (n=10 in 60 scenarios)	10.82 (9.43 – 12.21)	12.08 (10.69 – 13.47)	1.26	0.32	3.91	<0.001**	0.61	1.91
Participants' self-assessment (5-Likert scale)								
Domains								
Situation awareness (out of 5)	3.57 (3.42 – 3.72)	3.79 (3.63 – 3.94)	0.22	0.06	3.79	<0.001**	0.11	0.33
Decision making (out of 5)	3.53 (3.38 – 3.69)	3.83 (3.68 – 3.99)	0.30	0.06	5.26	<0.001**	0.19	0.41
Task management (out of 5)	3.52 (3.37 – 3.67)	3.84 (3.69 – 3.99)	0.32	0.06	5.47	<0.001**	0.21	0.44
Team working (out of 5)	3.83 (3.68 – 3.99)	4.05 (3.89 – 4.20)	0.21	0.06	3.80	<0.001**	0.10	0.33
Communication (out of 5)	3.75 (3.61 – 3.89)	3.98 (3.84 – 4.12)	0.23	0.06	4.05	<0.001^	0.12	0.34
Average overall (out of 5)	3.64 (3.50 – 3.78)	3.90 (3.76 – 4.04)	0.26	0.04	6.02	<0.001^	0.18	0.35

** Mean (95% CI) is the estimated marginal means with corresponding 95% confidence intervals of the linear mixed models*

*** A significant effect was found for the contents of other scenarios as compared to scenario 6 which was designated the reference category.*

^ A significant effect was found for the randomized schedules, and the contents of other scenarios as compared to scenario 6 which was designated the reference category.

The weighted Cohen kappa's level of agreement for each item ranged from fair to moderate (min–max=0.19–0.58) with less than 5% of items achieving an opposite rating (3 points scale: never vs consistently; Supplementary table S1).

A descriptive summary of the distribution of the mean Mayo scale scores per team shows that these scores improve when the Crisis Checklist App was used (*Figure 2A*). Similar favourable distributions in scenarios with the Crisis Checklist App were found for the percentages of omitted safety-critical steps, and the overall FoNTS matrix and participant self-assessment scores (*Figure 2*).

FoNTS matrix for non-technical skills

The mean score of the FoNTS matrix was higher for the scenarios with the checklists as compared to the scenarios without use of checklists (11.7 (95% CI =11.1 – 12.3) vs. 12.9 (95% CI =12.3–13.5) out of 16 points (Mean Difference=1.23 (95% CI =0.87 – 1.59); p<0.001; *Table 3*). All of the four domains within the FoNTS matrix showed significant improvements (*Table 3*).

An opposite rating (≥2 points on a 4 points scale) was scored in less than 8% of each item and the level of agreement of each item ranged from fair to moderate (weighted Cohen kappa min–max =0.28–0.52; Supplementary Table S2).

Participants self-assessment after each scenario

The self-assessment mean scores for each of the five domains ("situational awareness", "decision making", "task management", "teamwork", and "communication") demonstrated higher scores for the scenarios in which the checklists were used as compared to the scenarios without checklists (*Table 3*).

Performance of safety-critical steps

Of the 891 safety-critical 86 steps were omitted when the checklists were available as compared to 157 of 911 steps omitted when checklists were not available. This difference was significant and in favour of the scenarios in which the checklists were used (9.6% (95% CI =6.8–12.4) vs. 17.4% (95% CI =14.6–20.1); Mean Difference = -7.77 (95% CI = -10.73 to -4.81); p<0.001; *Table 3*).

More safety-critical steps were completed faster in scenarios with checklists as compared to scenarios without checklists (mean 52.2% (SD=22.6) vs. mean 44.0% (SD=23.0); mean difference = -8.2 (95% CI = -14.8 to -1.6); p=0.015). All safety-critical safety steps were completed by the teams more often in scenarios with the checklist (40 scenarios

vs 21 scenarios). The non-parametric survival analysis demonstrated that the time to complete all these steps was shorter in scenarios with checklists (log-rank test stratified by scenario $\chi^2 = 8.06$; $p = 0.005$, Supplementary Figure 1).

The LMMs demonstrates that the percentage of omitted safety-critical steps was lower if the person in charge of the group was a resident or nurse practitioner and not a senior medical student in the scenarios without checklists (14.9% (95% CI =11.6–18.2) vs. 22.6% (95% CI =17.7–27.5); Mean Difference = -7.69 (95% CI = -13.61 to -1.77); p=0.01; Supplementary Table S3). However, in the three scenarios when checklist were used the percentage of omitted safety-critical steps did not significantly differ between the groups of residents or nurse practitioners and senior students (7.93% (95% CI =4.6 – 11.3) vs. 13.4% (95% CI =8.4–18.4); Mean Difference = -5.46 (95% CI = -11.5 – 0.57); p=0.07; Supplementary Table S4).

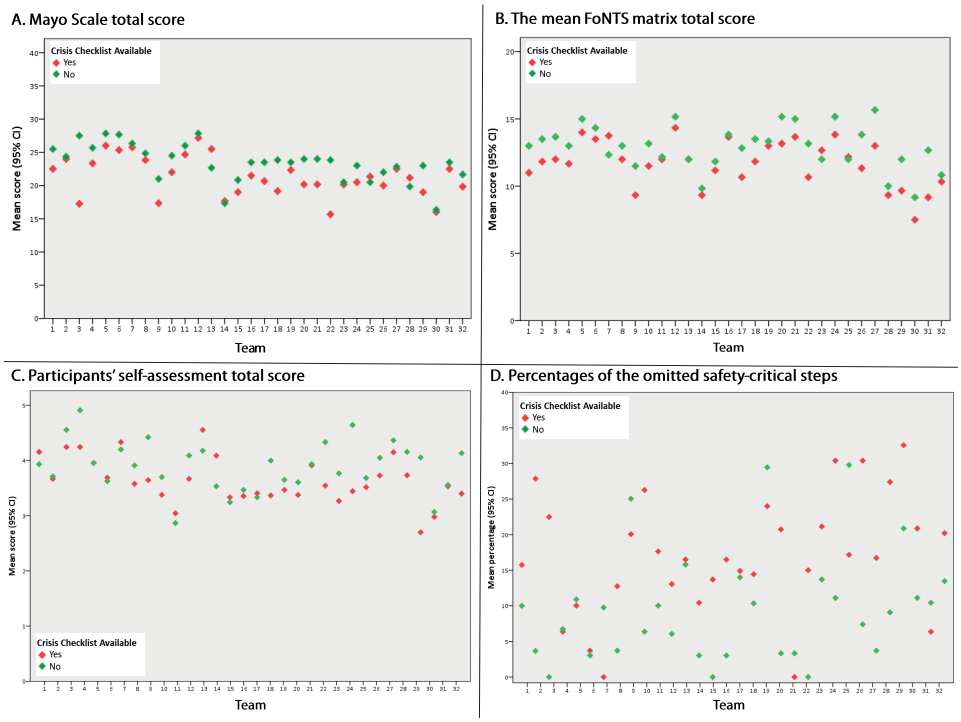


Figure 2. Descriptive summary with dot plots showing the distributions of the means for each team in scenarios with and without the checklist of the Mayo scale (A), FoNTS matrix (B), and Self-assessments scores (C), and the mean percentages of omitted safety-critical steps (D).

Learning effect

The mean scores of the Mayo scales and FoNTS matrix for all groups show a downward trend from the first until the third scenario when it was performed without the benefit of a checklist (*Figure 3*). A step change in these mean scores was seen after the implementation of the crisis checklists. These results remained unchanged with higher scores in the remaining two scenarios (*Figure 3*).

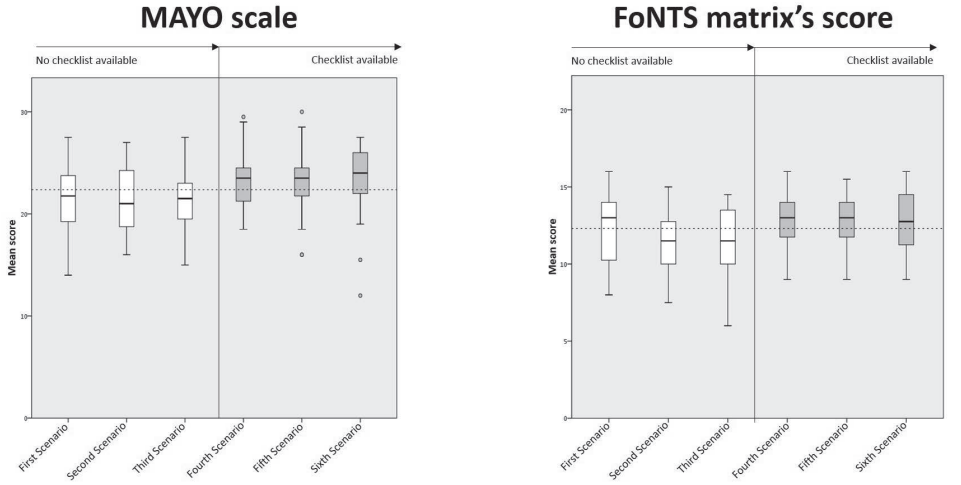


Figure 3. Learning effect

Mayo high performance resource scale's and FoNTS matrix's median scores for the scenarios, in order of performance, with and without the crisis checklist

DISCUSSION

This study provides important insights into the conceptual development and testing of clinical decision support tools for teams that respond to deteriorating patients with complex problems on general wards. Introducing the Crisis Checklist App in a multicentre simulation study of medical emergencies was associated with marked improvements in measured and self-reported teamwork. In addition, both the percentage of omitted predefined safety-critical steps and the time to complete these steps were lower in scenarios with the Crisis Checklist App. Non-technical skills assessment in all domains including "situational awareness", "decision making", "task management", and "communication" appeared to improve in both the observers and the participants' self-assessments.

The study design with the randomization of teams in two arms to complete matching scenarios in a different order was deliberately chosen to minimize selection bias and to compare the participants as their own controls. In addition, the analyses of the video recordings provided a comprehensive and independent assessment of non-technical skills and of acute care patient management based on a list of predefined safety-critical steps. However, also related to this study design is a learning effect bias since participants can perform better when the consecutive simulation are too similar. This bias might overestimate the effect of the Crisis Checklist App. But instead of a linear sequential increase, in this study a step change was seen for both the Mayo scale and FoNTS matrix scores directly after the introduction of the checklists which indicates that the improvement of these outcomes might be more attributable to the intervention.

Simulation-based trials investigating paper-based or computerized checklists have shown effective results in acute care settings.[23-27] The use of surgical crisis checklists by operating-room theatre teams resulted in improved crisis management and teamwork performance during simulated surgical-crisis scenarios.[23] The introduction of checklists in a simulated surgical day-care resulted in a 29% absolute reduction of omitted pre-defined key processes, but the non-technical skills did not improve.[27] This last result is in contrast to our findings which might be explained by the different setting, variation in the checklist design, or by the use of other instruments to assess the non-technical skills. Simulation studies investigating computerized checklists in acute care management have demonstrated similar reductions in omitted critical safety steps.[24-26, 28, 29]

More research is required to validate and determine the durability of these favourable effects in clinical practice. In addition, future development of the Crisis Checklist App should focus on the improvement of the user interface. This could include a built-in multi-language feature, or the ability to integrate the app into local electronic medical records (EMR) which might enhance usability by including the patient's EMR data.[24, 29] Notably, in this study the difference in the percentages of omitted safety-critical steps between teams that were led by a physician as compared to a medical student disappeared after the introduction of the Crisis Checklist App. This observation needs further investigation as it might hint to a key educational role for digital aids.

This study has several limitations. First, the unpredictable nature of clinical emergencies hampers interventional studies in deteriorating patients. Simulation provides an alternative and is valued for its ability to reproduce clinical conditions in a safe environment without endangering patients or clinicians.[30-33] The limitations of this study are therefore inherent to all simulation-based studies. We cannot rule out bias due to the method of team assignments, the performance of the facilitators, the script and content

of the scenarios, and the matching process of these scenarios. Second, weighted Cohen kappa's level of agreement between two observers for each item of the Mayo scale and FoNTS matrix ranged from fair to moderate, though an opposite rating was scored in less than 5% and 8% of the items respectively. Third, validity evidence for the tools to assess non-technical skills in simulated ward emergencies was not available. Finally, there are some challenges linked to the multicentre design of our study. Although the study protocol was centrally developed with the facilitators being trained at the same centre, differences are unavoidable in how simulation facilities manage and run their simulations which might have affected the study results.

CONCLUSION

Our findings suggest that a novel mobile crisis checklist application might be a valuable clinical decision support tool. We demonstrated improved teamwork performance and clinical decision making in a simulation-based study using an easily accessible checklist-based application for mobile devices to assist clinicians at the bedside. Further research is needed to determine the precise mechanisms, the role as an educational tool, and durability of these effects in clinical practice.

REFERENCES

1. Specialists DAoM. *Vision document*: Medical Specialist 2025. 2017
2. Clark AV, LoPresti CM, Smith TI. Trends in Inpatient Admission Comorbidity and Electronic Health Data: Implications for Resident Workload Intensity. *Journal of hospital medicine* 2018; **13**: 570-2
3. Steventon A, Deeny S, Friebel R, Gardner T, Thorlby R. Emergency hospital admissions in England: which may be avoidable and how? *The Health Foundation - Publications* 2018
4. Massey D, Chaboyer W, Anderson V. What factors influence ward nurses' recognition of and response to patient deterioration? An integrative review of the literature. *Nursing open* 2017; **4**: 6-23
5. Ludikhuize J, Dongelmans DA, Smorenburg SM, Gans-Langelaar M, de Jonge E, de Rooij SE. How nurses and physicians judge their own quality of care for deteriorating patients on medical wards: self-assessment of quality of care is suboptimal. *Critical care medicine* 2012; **40**: 2982-6
6. Thim T, Krarup NHV, Grove EL, Rohde CV, Løfgren B. Initial assessment and treatment with the Airway, Breathing, Circulation, Disability, Exposure (ABCDE) approach. *Int J Gen Med* 2012; **5**: 117-21
7. Smith D, Bowden T. Using the ABCDE approach to assess the deteriorating patient. *Nursing standard* 2017; **32**: 51-63
8. Bissell RA, Eslinger DG, Zimmerman L. The efficacy of advanced life support: a review of the literature. *Prehospital and disaster medicine* 1998; **13**: 77-87
9. Apple app store: https://apps.apple.com/gb/app/crisis-checklist/id987871235
10. Google Play: https://play.google.com/store/apps/details?id=uk.co.galactig.checklist
11. Subbe CP, Kellett J, Barach P, et al. Crisis checklists for in-hospital emergencies: expert consensus, simulation testing and recommendations for a template determined by a multi-institutional and multi-disciplinary learning collaborative. *BMC health services research* 2017; **17**: 334
12. Kreeftenberg HG, Pouwels S, van der Voort PHJ. Efficient organisation of intensive care units with a focus on quality: the non-physician provider. *Critical care* 2017; **21**: 118
13. Kellett J, Subbe, C. P., & Winsett, R. P. Recognizing and responding to the deteriorating patient. *Handbook of ICU Therapy* 2014: 221–34
14. de Bie AJR, Subbe CP, Bezemer R, et al. Differences in identification of patients' deterioration may hamper the success of clinical escalation protocols. *QJM : monthly journal of the Association of Physicians* 2019: **112**; 497–504
15. Malec JF, Torsher LC, Dunn WF, et al. The mayo high performance teamwork scale: reliability and validity for evaluating key crew resource management skills. *Simulation in healthcare* 2007; **2**: 4-10
16. Stocker M, Menadue L, Kakat S, et al. Reliability of team-based self-monitoring in critical events: a pilot study. *BMC emergency medicine* 2013; **13**: 22

17. Hamilton N, Freeman BD, Woodhouse J, Ridley C, Murray D, Klingensmith ME. Team behavior during trauma resuscitation: a simulation-based performance assessment. *Journal of graduate medical education* 2009; **1**: 253-9
18. Mellanby E.A. HM, Glavin R., Maran N. Development of a behavioural marker system for the non-technical skills of junior doctors in acute care 2015
19. JM Schraagen AS, M Smit, D van der Beek, J Van de Ven, P Barach. Improving methods for studying teamwork in cardiac surgery. *Quality & safety in health care* 2010; **19**: 1-6
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159-74
21. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007; **370**: 1453-7
22. Cheng A, Kessler D, Mackinnon R, et al. Reporting Guidelines for Health Care Simulation Research: Extensions to the CONSORT and STROBE Statements. *Simulation in healthcare* 2016; **11**: 238-48
23. Arriaga AF, Bader AM, Wong JM, et al. Simulation-based trial of surgical-crisis checklists. *The New England journal of medicine* 2013; **368**: 246-53
24. De Bie AJR, Nan S, Vermeulen LRE, et al. Intelligent dynamic clinical checklists improved checklist compliance in the intensive care unit. *British journal of anaesthesia* 2017; **119**: 231-8
25. Barwise A, Garcia-Arguello L, Dong Y, et al. Checklist for Early Recognition and Treatment of Acute Illness (CERTAIN): evolution of a content management system for point-of-care clinical decision support. *BMC medical informatics and decision making* 2016; **16**: 127
26. Vukoja M, Kashyap R, Gavrilovic S, Dong Y, Kilickaya O, Gajic O. Checklist for early recognition and treatment of acute illness: International collaboration to improve critical care practice. *World journal of critical care medicine* 2015; **4**: 55-61
27. Everett TC, Morgan PJ, Brydges R, et al. The impact of critical event checklists on medical management and teamwork during simulated crises in a surgical daycare facility. *Anaesthesia* 2017; **72**: 350-8
28. Thongprayoon C, Harrison AM, O'Horo JC, Berrios RA, Pickering BW, Herasevich V. The Effect of an Electronic Checklist on Critical Care Provider Workload, Errors, and Performance. *Journal of intensive care medicine* 2016; **31**: 205-12
29. Sevilla-Berrios R, O'Horo JC, Schmickl CN, et al. Prompting with electronic checklist improves clinician performance in medical emergencies: a high-fidelity simulation study. *International journal of emergency medicine* 2018; **11**: 26
30. Kothari LG, Shah K, Barach P. Simulation based medical education in graduate medical education training and assessment programs. *Progress in Pediatric Cardiology* 2017; **44**: 33-42
31. Cook DA, Hatala R, Brydges R, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *JAMA* 2011; **306**: 978-88
32. Lamé G, Dixon-Woods M. Using clinical simulation to study how to improve quality and safety in healthcare. *BMJ Simulation and Technology Enhanced Learning* 2018: bmjstel-2018-000370

33. Armenia S, Thangamathesvaran L, Caine AD, King N, Kunac A, Merchant AM. The Role of High-Fidelity Team-Based Simulation in Acute Care Settings: A Systematic Review. *Surgery journal* 2018; **4**: e136-e51

APPENDIX 1: STUDY PROTOCOL

Study protocol available on request.

APPENDIX 2: SUPPLEMENTARY DATA.

Table S1. Inter-observer agreement for the Mayo scale.

Mayo scale item	Scorable items	Weighted Kappa	Weighted Kappa – 95% CI	Percentage of answers with a certain level of difference between the answers.			
				0 points	1 point	2 points	>2 points
Mayo 1	3	0.54	0.41 – 0.66	81	18	1	n.a.
Mayo 2	3	0.54	0.42 – 0.66	77	22	1	n.a.
Mayo 3	3	0.47	0.33 – 0.60	77	23	0	n.a.
Mayo 4	3	0.28	0.13 – 0.42	67	31	2	n.a.
Mayo 5	3	0.29	0.15 – 0.43	68	30	2	n.a.
Mayo 6	3	0.34	0.22 – 0.45	60	38	2	n.a.
Mayo 7	3	0.44	0.33 – 0.54	57	40	3	n.a.
Mayo 8	3	0.33	0.19 – 0.47	73	26	1	n.a.
Mayo 9	3	0.19	-0.02 – 0.40	67	28	5	n.a.
Mayo 10	3	0.35	0.20 – 0.51	89	10	1	n.a.
Mayo 11	3	0.41	0.26 – 0.56	69	31	0	n.a.
Mayo 12	3	0.58	0.43 – 0.72	79	21	0	n.a.
Mayo 13	3	0.33	0.17 – 0.50	68	29	3	n.a.
Mayo 14	3	0.27	0.10 – 0.43	70	27	3	n.a.
Mayo 15	3	0.25	0.06 – 0.45	78	22	0	n.a.
Mayo 16	3	0.40	0.25 – 0.55	68	31	1	n.a.

n.a. = not applicable

Table S2. Inter-observer agreement for the FoNTS matrix.

FoNTS items	Scorable items	Weighted Kappa	Weighted Kappa – 95% CI	Percentage of answer with a certain level of difference between the answers.			
				0 points	1 point	2 points	>2 points
FoNTS 1.1	4	0.36	0.23 – 0.49	55	41	3	1
FoNTS 1.2	4	0.45	0.32 – 0.58	63	30	7	0
FoNTS 1.3	4	0.38	0.25 – 0.51	58	38	4	0
FoNTS 1 sum Linear	12	0.43	0.33 – 0.52	32	35	26	7
FoNTS 2.1	4	0.41	0.28 – 0.54	57	42	1	0
FoNTS 2.2	4	0.41	0.29 – 0.53	54	42	4	0
FoNTS 2.3	4	0.45	0.33 – 0.57	59	37	4	0
FoNTS 2 sum Linear	12	0.43	0.33 – 0.52	26	39	17	18
FoNTS 3.1	4	0.28	0.15 – 0.41	51	43	5	1
FoNTS 3.2	4	0.36	0.23 – 0.49	58	37	5	0
FoNTS 3.3	4	0.43	0.30 – 0.57	64	23	3	0
FoNTS 3.4	4	0.29	0.16 – 0.42	51	44	5	0
FoNTS 3 sum Linear	16	0.39	0.29 – 0.49	27	31	23	19
FoNTS 4.1	4	0.30	0.17 – 0.44	52	47	1	0
FoNTS 4.2	4	0.31	0.17 – 0.45	53	41	5	1
FoNTS 4.3	4	0.35	0.22 – 0.49	54	44	2	0
FoNTS 4 sum Linear	12	0.37	0.27 – 0.48	28	35	27	10
FoNTS 1 overall	4	0.50	0.40 – 0.61	66	32	2	0
FoNTS 2 overall	4	0.41	0.31 – 0.51	57	40	3	0
FoNTS 3 overall	4	0.52	0.42 – 0.63	69	28	3	0
FoNTS 4 overall	4	0.45	0.33 – 0.56	65	33	2	0
FoNTS overall Linear	16	0.51	0.43 – 0.59	36	30	22	12

Table S3. Percentages of omitted safety-critical steps overall and per subgroup comparisons based on the scenarios with and without checklists

	Scenarios without checklists Mean (95% CI)*	Scenarios with checklists Mean (95% CI)*	Mean Difference	Standard Error	Paired t-test	p-value	95% Confidence Interval	
							Lower Bound	Upper Bound
Omitted safety-critical steps (%)	17.36 (14.62 – 20.11)	9.60 (6.83 – 12.36)	-7.77	1.50	-5.19	<0.001**	-10.73	-4.81
Omitted safety-critical steps based on the person in the lead								
Residents (n=19 in 111 scenarios)	13.84 (10.45 – 17.23)	8.03 (4.63 – 11.44)	-5.80	1.59	-3.64	<0.001**	-8.97	-2.64
Nurse practitioners (n=3 in 17 scenarios)	21.93 (12.37 – 31.50)	5.54 (-4.02 – 15.10)	-16.40	6.07	-2.70	0.022**	-29.92	-2.87
Senior students (n=10 in 60 scenarios)	22.21 (17.25 – 27.16)	13.85 (8.79 – 18.90)	-8.36	3.04	-2.75	0.009**	-14.49	-2.23

** Mean (95% CI) is the estimated marginal means with corresponding 95% confidence intervals of the linear mixed models*

*** A significant effect was found for the contents of other scenarios as compared to scenario 6 which was designated the reference category.*

Table S4. Percentages of omitted safety-critical steps comparison based on the person lead the scenarios with and without checklists

	Person in lead: Residents** Mean (95% CI)*	Person in lead: Senior students Mean (95% CI)*	Mean Difference	Standard Error	Paired t-test	p-value	95% Confidence Interval	
							Lower Bound	Upper Bound
Scenarios without checklists								
Omitted safety-critical steps (%)	14.91 (11.60 – 18.21)	22.60 (17.72 – 27.47)	-7.69	2.89	-2.66	0.01	-13.61	-1.77
Scenarios with checklists:	**Residents****	**Senior students**						
Omitted safety-critical steps (%)	7.93 (4.58 – 11.28)	13.39 (8.40 – 18.38)	-5.46	2.95	-1.85	0.07	-11.49	0.57

** Mean (95% CI) is the estimated marginal means with corresponding 95% confidence intervals of the linear mixed models*

*** Nurse practitioners had the same job description and responsibilities as residents in their daily clinical practice. Therefore, the residents and nurse practitioners were grouped together.*

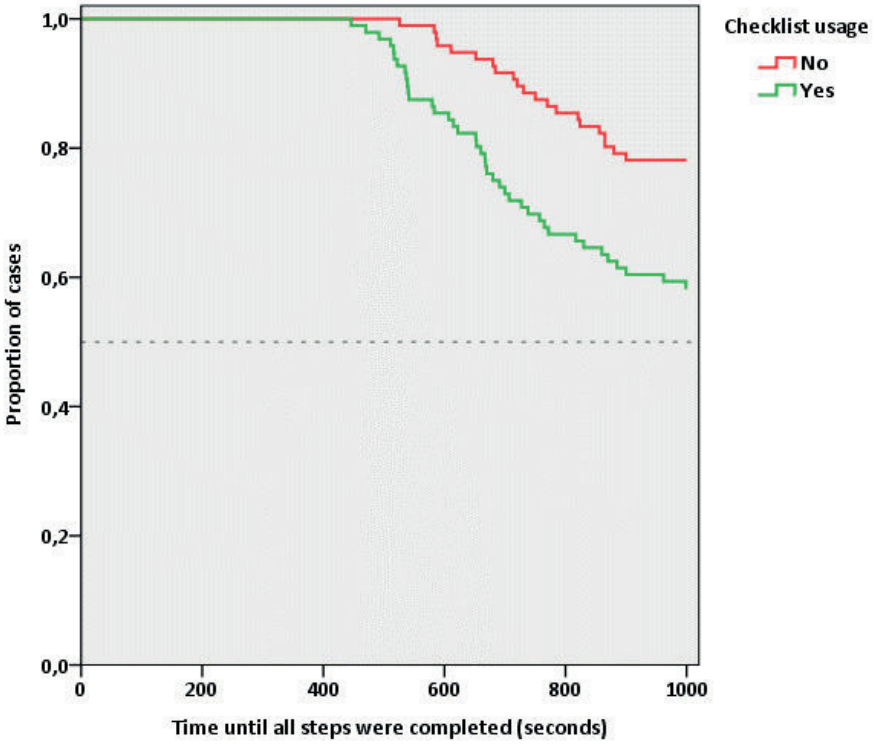
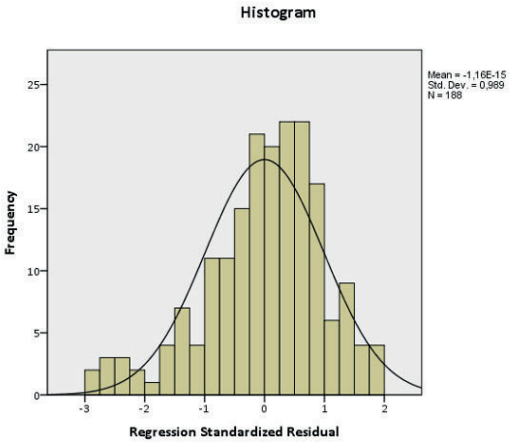
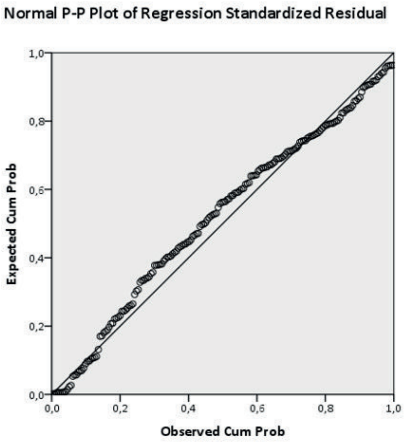


Figure S1: Kaplan-meier curve of the non-parametric test analysing the time until all scenarios were completed (log-rank test stratified by scenario $\chi 2 = 8.06$; $p = 0.005$).

APPENDIX SUPPLEMENTARY DATA: DISTRIBUTIONS OF RESIDUALS.

1. MAYO scale:

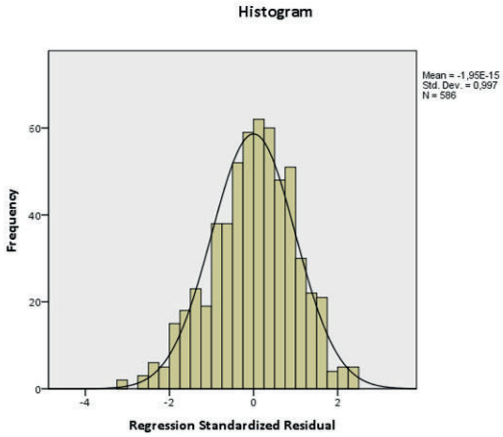


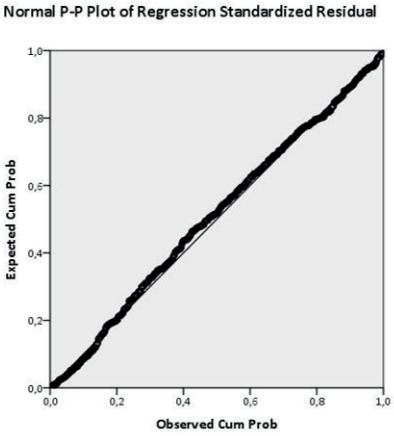


Regression for checking multicollinearity (variance inflation factor <5 (VIF))

	Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model	**B**	**Std. Error**	**Beta**	**t-test**	**Sig.**	**Tolerance**	**VIF**
1 (Constant)	20,41	1,04		19,56	<0,01		
Checklist used or not used	1,98	0,41	0,29	4,85	<0,01	1,000	**1,000**
Scenario type (1-6)	-0,14	0,12	-0,07	-1,14	0,26	0,998	**1,002**
Schedule A vs B	1,43	0,41	0,21	3,50	<0,01	0,999	**1,001**
Team	-0,16	0,02	-0,44	-7,24	<0,01	0,998	**1,002**

2. FoNTS matrix overall score:

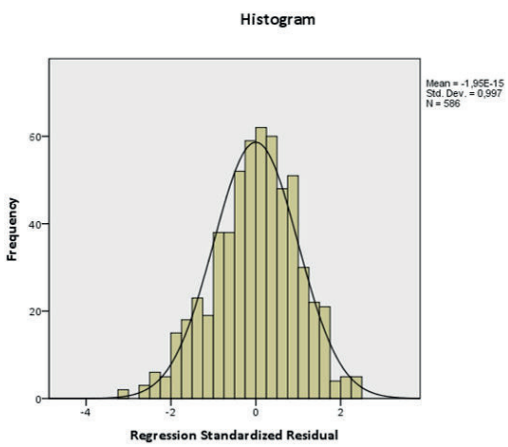


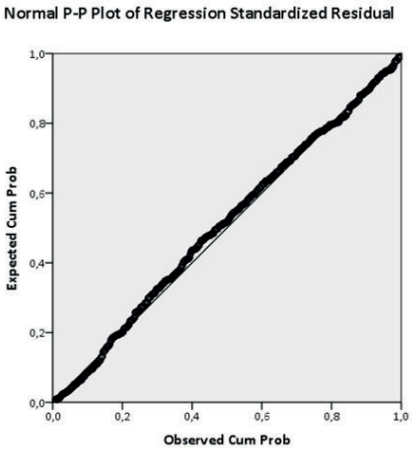


Regression for checking multicollinearity (variance inflation factor <5 (VIF))

	Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model	**B**	**Std. Error**	**Beta**	**t-test**	**Sig.**	**Tolerance**	**VIF**
1 (Constant)	10,61	6,69		15,32	<0,01		
Checklist used or not used	1,22	0,27	0,30	4,49	<0,01	1,000	1,000
Scenario type (1-6)	-0,10	0,08	-0,09	-1,29	0,20	0,998	1,002
Schedule A vs B	0,72	0,27	0,18	2,64	<0,01	0,999	1,001
Team	-0,05	0,02	-0,23	-3,50	<0,01	0,998	1,002

3. Participants' self-assessment survey: overall score

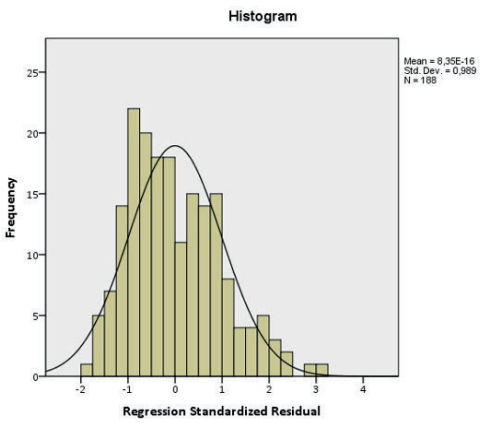


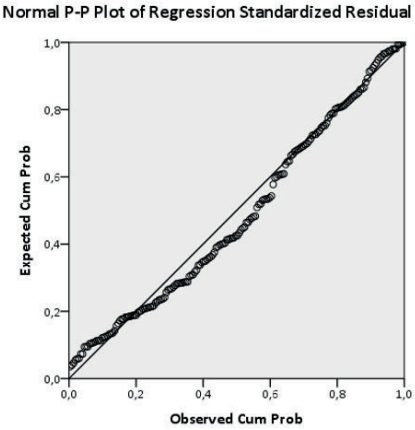


Regression for checking multicollinearity (variance inflation factor <5 (VIF))

	Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model	**B**	**Std. Error**	**Beta**	**t-test**	**Sig.**	**Tolerance**	**VIF**
1 (Constant)	3,30	0,13		26,10	0,00		
Checklist used or not used	0,26	0,05	0,20	5,19	0,00	0,999	1,001
Scenario type (1-6)	-0,05	0,02	-0,12	-3,04	0,00	0,999	1,001
Schedule A vs B	0,30	0,05	0,23	6,03	0,00	0,997	1,003
Team	-0,01	0,00	-0,179	-4,63	0,00	0,998	1,002

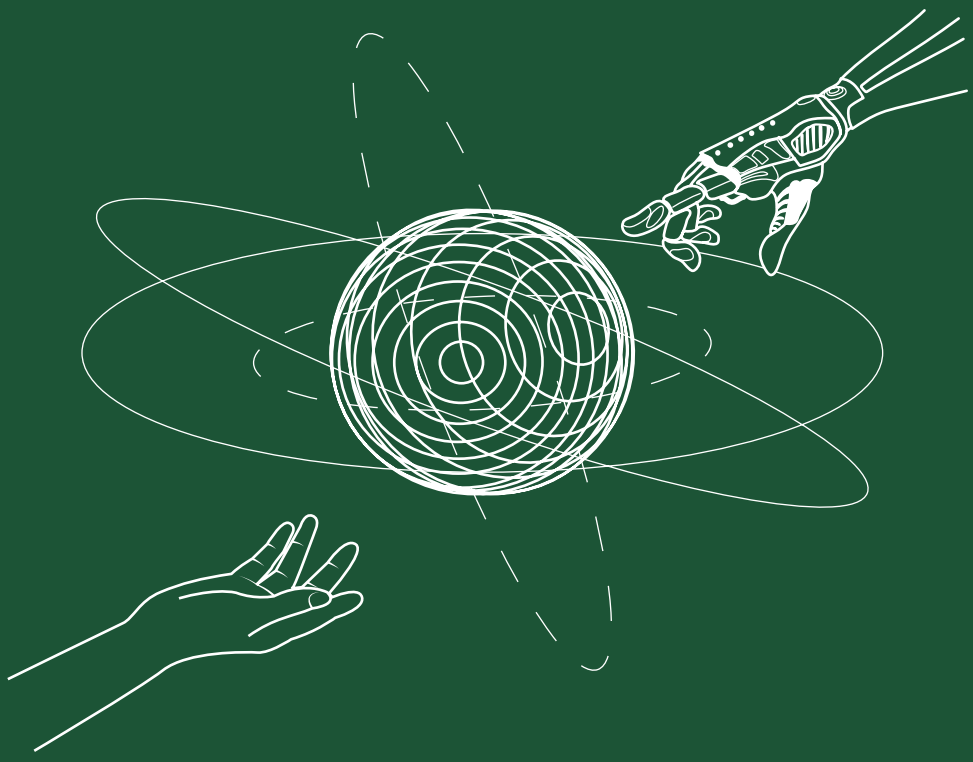
4. Percentages of omitted safety-critical steps





Regression for checking multicollinearity (variance inflation factor <5 (VIF))

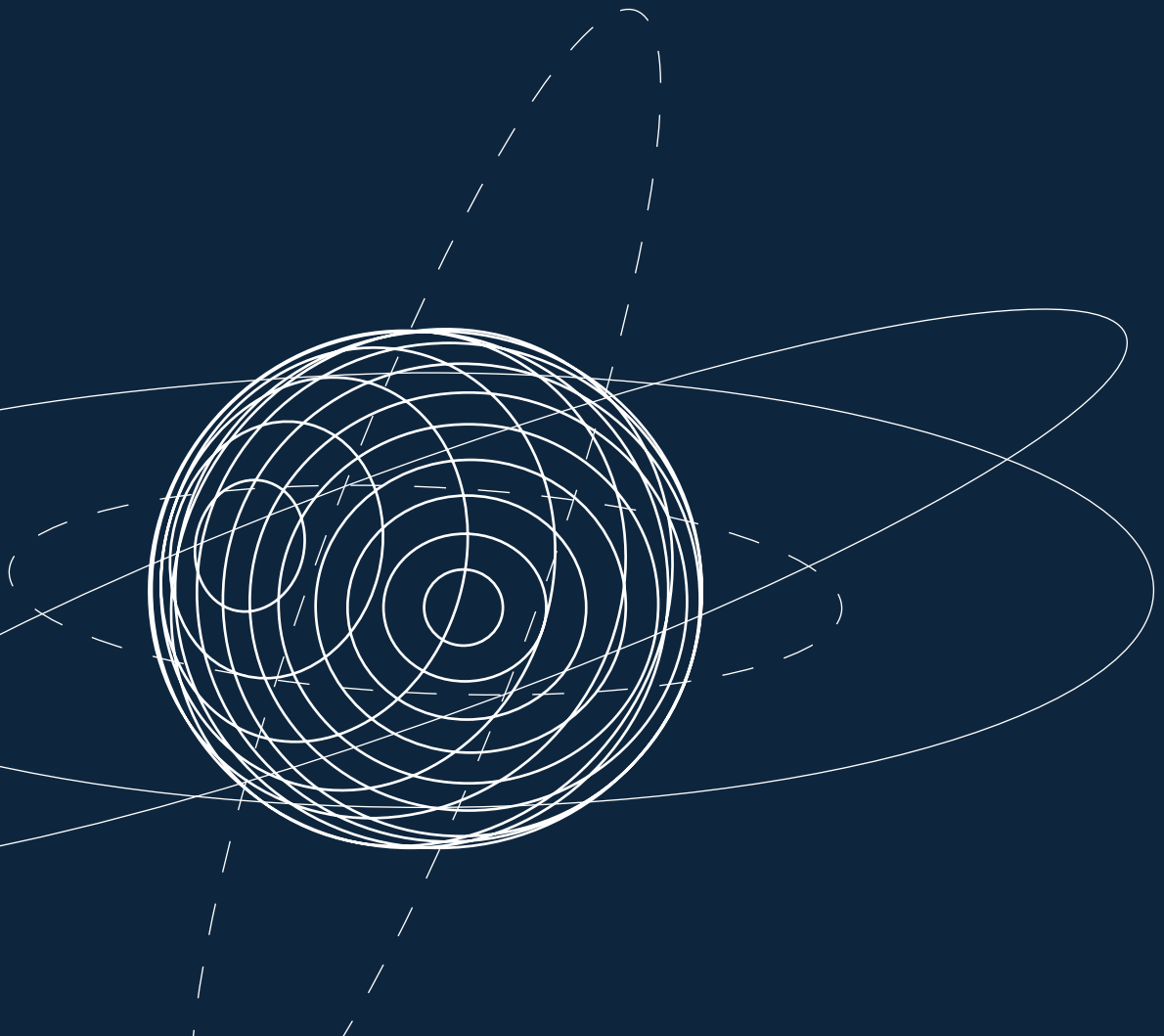
	Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model	**B**	**Std. Error**	**Beta**	**t-test**	**Sig.**	**Tolerance**	**VIF**
1 (Constant)	18,90	4,47		4,23	<0,01		
Checklist used or not used	-7,59	1,76	-0,30	-4,32	<0,01	1,000	1,000
Scenario type (1-6)	0,86	0,53	0,11	1,66	0,10	0,999	1,001
Schedule A vs B	-0,74	1,76	-0,04	-0,42	0,67	0,997	1,003
Team	0,24	0,10	0,17	2,52	0,01	0,997	1,003



PART 3

A CDS system that is based on a simple model which can gather some data automatically

3.1



Implementation of an automated early warning scoring system in a surgical ward: practical use and effects on patient outcomes

A.J.R. De Bie, E.H.J. Mestrom, M. van de Steeg, M. Driessen, L. Atallah, R. Bezemer, R.A. Bouwman, H.H.M. Korsten

PLOS ONE 2019

ABSTRACT

Introduction

Early warning scores (EWS) are being increasingly embedded in hospitals over the world due to their promise to reduce adverse events and improve the outcomes of clinical patients. The aim of this study was to evaluate the clinical use of an automated modified EWS (MEWS) for patients after surgery.

Methods

This study conducted retrospective before-and-after comparative analysis of non-automated and automated MEWS for patients admitted to the surgical high-dependency unit in a tertiary hospital. Operational outcomes included number of recorded assessments of the individual MEWS elements, number of complete MEWS assessments, as well as adherence rate to related protocols. Clinical outcomes included hospital length of stay, in-hospital and 28-day mortality, and ICU readmission rate.

Results

Recordings in the electronic medical record from the control period contained 7929 assessments of MEWS elements and were performed in 320 patients. Recordings from the intervention period contained 8781 assessments of MEWS elements in 273 patients, of which 3418 were performed with the automated EWS system. During the control period, 199 (2.5%) complete MEWS were recorded versus 3991 (45.5%) during intervention period. With the automated MEWS systems, the percentage of missing assessments and the time until the next assessment for patients with a MEWS of ≥2 decreased significantly. The protocol adherence improved from 1.1% during the control period to 25.4% when the automated MEWS system was involved. There were no significant differences in clinical outcomes.

Conclusion

Implementation of an automated EWS system on a surgical high dependency unit improves the number of complete MEWS assessments, registered vital signs, and adherence to the EWS hospital protocol. However, this positive effect did not translate into a significant decrease in mortality, hospital length of stay, or ICU readmissions. Future research and development on automated EWS systems should focus on data management and technology interoperability.

INTRODUCTION

Automated early warning score (EWS) systems are increasingly embedded in clinical practice to improve registration and awareness of vital signs and enhance rapid response teams (RRT) notifications.[1] The impact on clinical outcomes of these systems remains uncertain in various populations, for example in the high-risk surgical population, which is the focus of this study.

Complications that are frequently encountered on the general ward can lead to major adverse events such as unplanned intensive care unit (ICU) admissions, cardiorespiratory arrest, and mortality.[2,3] The EWS has been developed as an objective bedside tool to help clinicians identify patients at risk of adverse events.[4,5] Since its introduction in the late nineties, studies have shown varied results on the predictive value of EWS as well as its implications on clinical outcomes.[4, 6-9] Consequently, newer versions of EWS have been developed to improve clinical relevance, such as the modified early warning score (MEWS).[10] Digital automated systems have been introduced to assist clinicians in completing the EWS assessment. Several studies emphasized that these systems provide a faster completion of EWS with increased accuracy. This is important since complete and accurate registration on a regular base is essential for the effectiveness of EWS, especially when upcoming assisting medical technologies, such as clinical decision support tools, rely on these data.[1, 11-16]

In addition, many studies have shown that automated systems reduce mortality and length of hospital stay. These studies also showed an improvement in the survival of patients treated by RRTs after EWS-triggered notification.[1, 11, 17, 18] On the other hand, the study of Dawes et al. showed no significant improvement in mortality among 3184 patients admitted to an acute medical unit after an electronic alerting physiological scoring system was introduced.[19]

The automation of MEWS systems may be more effective for specific subgroups of patients such as high-risk surgical patients on the ward since postoperative complications develop more often and documentation on vital signs is known to be lacking there.[20, 21]

The aim of this retrospective before-and-after cohort study is to investigate whether an automated MEWS system on a surgical high dependency unit (HDU) had a positive effect on clinical practice, in terms of improved documentation of vital signs and complete EWS assessments, and EWS protocol adherence. Secondary aim of the study was to evaluate the impact on clinical outcomes, such as mortality, length of stay, and ICU readmissions.

METHODS

Study design

This is a retrospective before-and-after comparative analysis of clinical practice and outcomes for patients admitted to the eight-bed surgical HDU in the Catharina Hospital, a tertiary teaching hospital in Eindhoven, the Netherlands. This unit functions as a step-down unit between the ICU and the regular surgical ward in the postoperative phase after major elective and acute surgeries. These surgeries include major gastro-intestinal, oncologic, and vascular surgeries, such as pylorus preserving pancreaticoduodenectomy, open vascular aortic surgery and hyperthermic intraperitoneal chemotherapy.

Inclusion criteria were patients admitted for step-down care after surgery, who required an ICU admission for postoperative hemodynamic surveillance. Exclusion criteria were patients under 18 years old, ICU admission for other reasons than hemodynamic surveillance, such as electrolyte surveillance after thyroidectomy, and patients who had a second admission on the HDU after a hospital discharge during the study period or who were admitted for other reasons than step-down care.

The study period consisted of two 15 month phases: the control phase (January 2012 until March 2013) and the intervention phase (June 2013 until August 2015). Data from a three month period between the phases was omitted to account for any influences of the training period directly after implementation of the automated MEWS system, the Philips IntelliVue Guardian Solution (Guardian®). The study was approved by the Medical research Ethics Committees United (MEC-U; study ID: non-WMO 2015–87). The study was classified as non-WMO by the MEC-U based on the retrospective design. Therefore obtaining an informed consent was not deemed necessary as it conforms to the Dutch Agreement on Medical Treatment Act. All data were analysed anonymously.

Control period

In correspondence with common daily practice, the MEWS was used during the control period. This MEWS was introduced in 2011 in the hospital during a national study by Ludikhuize et al (*Table 1*).[17] MEWS parameters are heart rate, oxygen saturation (SpO2), respiratory rate, non-invasive blood pressure, temperature, AVPU scale, and 24-hour urine production. According to the MEWS hospital protocol, these parameters were assessed at bedside within certain timeframes based on the previous MEWS (*Table 2*). All measured parameters were manually recorded in the electronic medical record (EMR). Nurses could manually calculate the MEWS using cards containing the MEWS algorithm. They were instructed to alert the physician on call in case of a MEWS ≥3 for assessment

and potential treatment (*Table 2*). Nurses could also escalate to call the RRT of the ICU directly if the physician on call were not available or if initiated treatments after a certain MEWS did not lead to any improvement.

Table 1. Modified early warning score system

Score	3	2	1	0	1	2	3
Heart rate (beats/min)		<40	40-50	51-100	101-110	111-130	>130
Systolic blood pressure (mm Hg)	<70	70-80	81-100	101-200		>200	
Respiratory rate (breaths/min)		<9		9-14	15-20	21-30	>30
Temperature (°C)		<35.1	35.1-36.5	36.6-37.5	>37.5		
Level of consciousness				A (Alert)	V (Voice responsive)	P (Pain responsive)	U (Unconscious)

Worried about patient's condition: 1 point
Urine production below 75ml during previous 4hrs: 1 point
Oxygen saturation below 90% despite adequate oxygen therapy: 3 points

Table 2. Response system for modified early warning score

MEWS	Time till next MEWS assessment
0	Next shift (within 24 hours)
1	Within 8 hours
2	Within 4 hours
3	Within 1 hour; Consult responsible physician
≥ 4	Within 1 hour; Consult responsible physician and consider to consult RRT

Intervention period

The electronic EWS system Philips IntelliVue Guardian Solution (Guardian®) was implemented on the HDU over a three-month period between the control and intervention phase. This system facilitated the acquisition of vital signs and the completion of MEWS to provide automated clinical decision support and awareness to the nursing staff. The device consisted of two spot-check monitors which were taken to the bedside of the patients to measure respiratory rate, non-invasive blood pressure, heart rate and SpO_2. Urine output, level of consciousness (AVPU scale), temperature, and the nurse's level of concern were manually entered in the Guardian® software. The device

calculated the MEWS values and showed them on the screen of the device as well as on a monitor at the central nurse station. In addition to the MEWS value, a short advice was displayed on the screen for further monitoring, such as the recommended time until the next assessment, or recommended actions, such as alerting a physician or the RRT. In addition, the monitor at the central nurse station also displayed these features. Every spot-check observation was stored in a database, which was not connected to the electronic medical record (CS-EZIS test, Chipsoft BV, Amsterdam, The Netherlands). Nurses were required to copy the measured vital parameters and MEWS into the EMR. All nurses received appropriate training on data collection before implementation of the Guardian® system. All parameters needed to be recorded within a 15-minute timeframe to be considered in a single MEWS measurement. During the intervention period, the conventional methods used during the control period were still available to collect MEWS parameters and record them in the EMR, and although nurses were trained to use the automated system, they were free to work according to their preference.

Data collection and outcomes

The following data were retrieved from the EMR or the Guardian database during the periods: patient characteristics, vital signs and all other elements from the MEWS, and outcomes such as ICU re-admission, mortality, and length of stay. APACHE II, APACHE IV and SAPS II scores were collected during the first postoperative admission at the ICU or in the event of an ICU readmission. The results of this study were divided into two categories: operational outcomes (primary outcomes) and clinical outcomes (secondary outcomes).

Primary outcomes

Operational outcomes under study include the practical clinical use of the automated EWS system, such as the number of documented MEWS elements and calculated MEWS values recorded in the EMR. In addition, the percentage of complete MEWS, daily patterns of the assessments, and the time interval between assessments were calculated. Time interval between assessments was also used to see if the next assessment was performed conform MEWS hospital protocol (*Table 1*). Measurements with a time interval of less than 15 minutes were considered a single assessment.

Secondary outcomes

The clinical outcomes analysed include the impact of the use of the automated MEWS system on the length of stay in the hospital, in-hospital and 28-day mortality, and ICU readmission rate. An ICU readmission was defined as a transfer to a higher level of care, such as ICU, medium care, or cardiac care unit.

Sub-analysis was performed for readmitted patients at the ICU; APACHE II, APACHE IV, and SAPS II scores on admission to the ICU were compared between the control and intervention group. Other outcomes for this subgroup analysis were the length of stay at the ICU and ICU mortality. In addition, a similar subgroup-analysis was performed for different age groups (≤49 years, 50–69 years and ≥70 years) or if certain patient characteristics differed significantly between the control and intervention period.

Statistical analysis

All statistical analyses were performed using IBM SPSS Statistics version 23 (IBM Corp, Armonk, New York, released 2015).[22] Normality of the data sets was tested using a Kolmogorov-Smirnov test. Continuous data are presented as means and standard deviations or medians and interquartile data based on the distribution of the data. Categorical data are presented as proportions or percentages. The Mann-Whitney test was used to test for differences in continuous variables with non-normal distributions and the chi-square test Fisher's exact test was used to test for differences in categorical groups. A value of $p<0.05$ was considered statistically significant.

RESULTS

A total of 594 patients were included for analysis, 320 patients for the control group and 274 patients for the intervention group. Both groups were comparable in terms of age, gender, and APACHE II and -IV scores at the time of first ICU admittance postoperatively (*Table 3*). The control group consisted of significantly less patients undergoing oncologic abdominal surgery (69.7 versus 80.3%; p=0.01) and more patients undergoing aortic surgery (20.0 versus 12.4%; p=0.01).

Operational outcomes

During the control period, a total of 7929 records of one or more elements of the MEWS were retrieved from the EMR. During the intervention period the total number of recorded assessments was 8781. Of these 8781 assessments, 3418 (39%) were recorded with the automated MEWS system. The other 5363 (61.1%) assessments were recorded with conventional monitoring systems. Results are shown in *Figure 1*.

Table 3. Baseline characteristics of all included patients (A) and patients readmitted to ICU (B)

A. Patients overall	**Control group**	**Intervention group**	**p-value**
Number of patients	320	273	
Age in years, median (IQR)	67 (15)	67 (16)	0.32
Male gender, number (%)	204 (63.8)	173 (63.4)	0.92
Unplanned (acute) surgery (%)	57 (18.1)	37 (13.6)	0.16
Type of surgery			0.01
Abdominal, oncology (%)	223 (69.7)	219 (80.2)	
Abdominal, benign (%)	33 (10.3)	20 (7.3)	
Vascular, aortic (%)	64 (20.0)	34 (12.5)	
APACHE II first ICU admission, median (IQR)	14 (6)	14 (6)	0.16
APACHE IV first ICU admission, median (IQR)	37 (15)	38 (18)	0.38
SAPS II first ICU admission, median (IQR)	31 (15)	31 (17)	0.47
B. Patients readmitted ICU	**Control group**	**Intervention group**	**p-value**
Number of patients	43	29	
Age in years, median (IQR)	65 (19)	67 (10)	0.30
Male gender, number (%)	33 (76.7)	23 (79.3)	0.80
Unplanned (acute) surgery (%)	8 (18.6)	4 (13.8)	0.83
Type of surgery			0.17
Abdominal, oncology (%)	29 (64.4)	25 (86.2)	
Abdominal, benign (%)	8 ((18.6)	2 (6.9)	
Vascular, aortic (%)	6 (14.0)	2 (6.9)	
APACHE II, median (IQR)			
First ICU admission	14 (4)	16 (6)	0.02
ICU readmission	20 (11)	19 (8)	0.35
APACHE IV, median (IQR)			
First ICU admission	37 (13)	43 (30)	0.08
ICU readmission	52 (50)	52 (20)	0.55
SAPS II, median (IQR)			
First ICU admission	29 (9)	33 (23)	0.18
ICU readmission	40 (24)	38 (21)	0.46
ICU interventions, number (%)			
Arterial line	32 (74.4)	21 (72.4)	0.85
Vasopressor use	16 (37.2)	11 (37.9)	0.95
Mechanical ventilation	29 (67.4)	14 (48.3)	0.10

The adherence to MEWS hospital protocol improved when the automated MEWS system was involved, from 1.1% (88 of 7929 assessments) in the control group to 25.4% (2237 of 8781 assessments) in the intervention group. Within the intervention group, the adherence to MEWS hospital protocol improved from 10.8% (599 of 5363 assessments) when a conventional method was used to 47.9% (1638 of 3418 assessments) when the automated MEWS system was used (*Figure 1*).

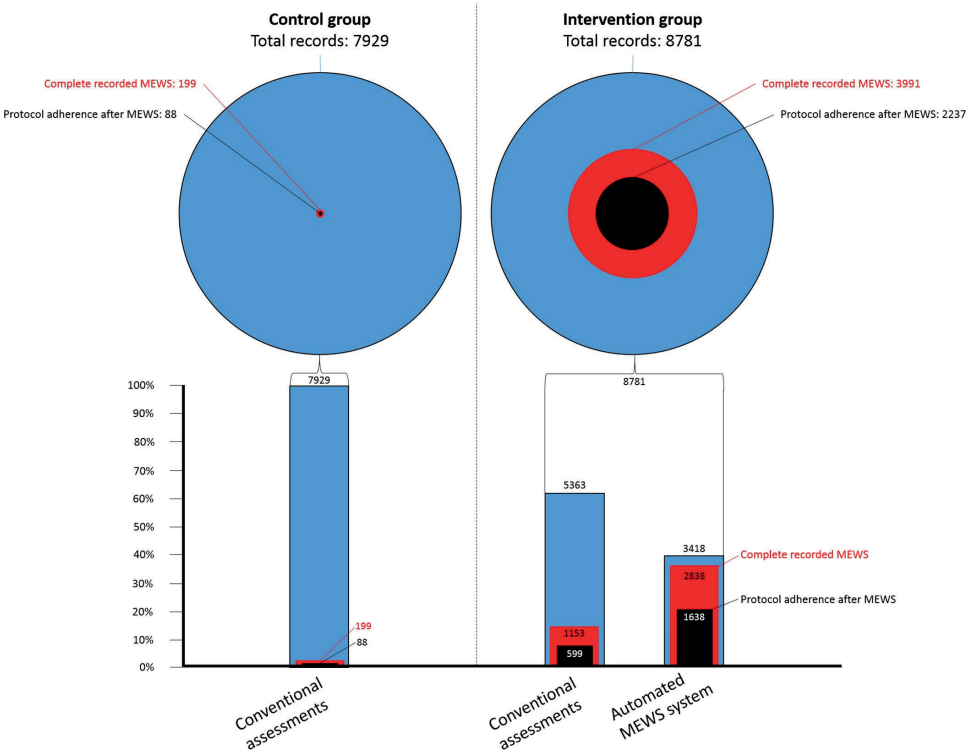


Figure 1. MEWS protocol adherence

The number of complete recorded MEWS and protocol adherence according to these MEWS for the control period and the intervention period, subdivided in conventional and automated MEWS assessments during the intervention period.

The implementation of the automated MEWS system resulted in significantly more completed MEWS containing all the MEWS elements from 199 (2.5%) complete assessments in the control group to 3991 (45.5%) complete assessments in the intervention group ($p<0.001$).

The number and percentage of missing elements in the MEWS for each documented record is shown in *Figure 2*. The most pronounced difference was observed for respiratory rate (96% versus 3%) and level of consciousness (100% versus 3%). The daily pattern is shown in *Figure* 3 and was comparable in both groups with three peaks during 24 hours corresponding to the daily nursing rounds. The median time until the next assessment was significantly shorter when a MEWS of 2 and higher was measured with the automated MEWS system compared to the conventional method between both groups and within the intervention group (*Table 4*).

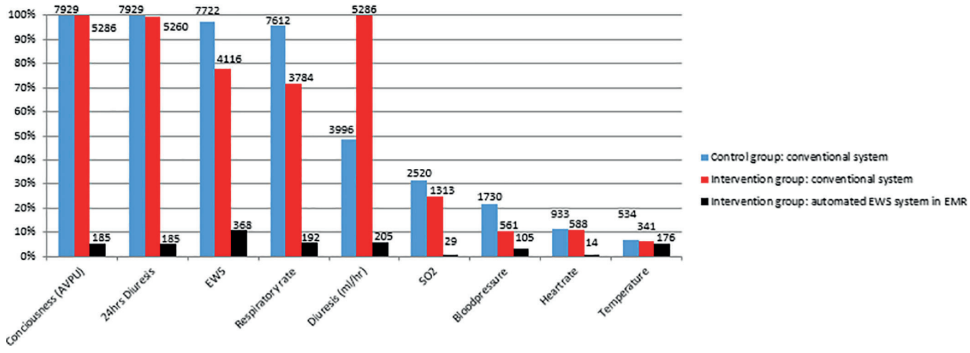


Figure 2. Percentages and absolute numbers of missing MEWS elements assessments

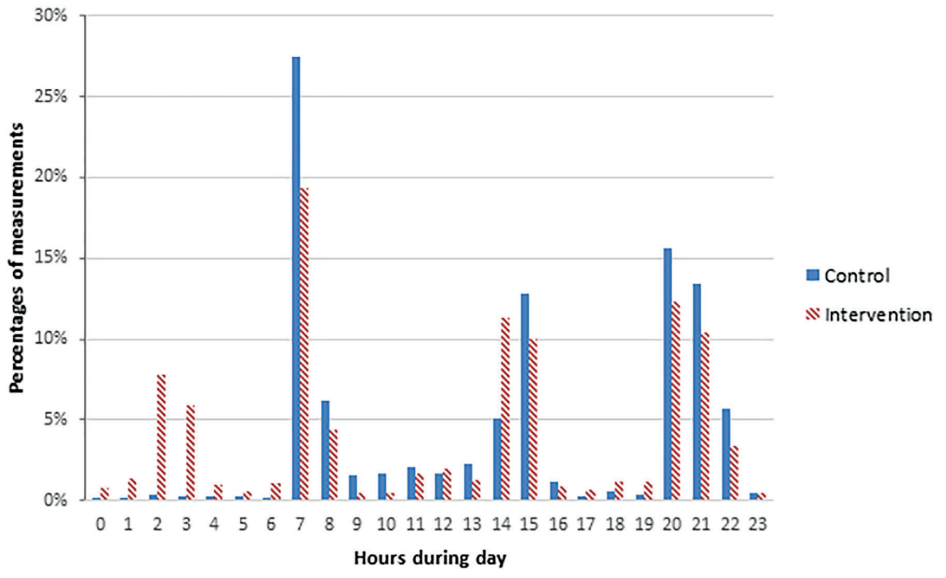


Figure 3. The 24-hour pattern of MEWS assessments. Pattern for control and intervention group.

Table 4. Median time in hours until next assessment.

	Control period Conventional system		Intervention period Conventional & Automated EWS system			
MEWS	**N**	**Median hours (IQR)**	**N**	**Median hours (IQR)**	**p value**	**Z-score**
0	33	7.40 (4.1)	668	6.06 (3.49)	0.011	-2.539
1	59	6.02 (5.46)	1370	5.86 (3.22)	0.35	-0.931
2	42	6.66 (4.32)	1144	5.68 (3.42)	0.011	-2.539
≥3	65	5.10 (6.79)	809	4.62 (3.21)	0.019	-2.338
	Intervention period Conventional system		**Intervention period Automated EWS system**			
MEWS	**N**	**Median hours (IQR)**	**N**	**Median hours (IQR)**	**p value**	**Z-score**
0	186	6.01 (4.72)	482	6.09 (3.09)	0.36	-0.922
1	388	5.78 (4.00)	982	5.95 (3.01)	0.36	-0.914
2	328	6.08 (3.72)	816	5.55 (3.44)	0.002	-3.037
≥3	251	5.06 (4.60)	558	4.29 (5.30)	<0.001	-3.997
	Control period Conventional system		**Intervention period Automated EWS system**			
MEWS	**N**	**Median hours (IQR)**	**N**	**Median hours (IQR)**	**p value**	**Z-score**
0	33	7.40 (4.11)	482	6.09 (3.09)	0.014	-2.467
1	59	6.02 (5.46)	982	5.95 (3.01)	0.4	-0.838
2	42	6.66 (4.32)	816	5.55 (3.44)	0.004	-2.879
≥3	65	5.10 (6.79)	558	4.29 (5.30)	0.003	-2.992

Clinical outcomes

There were no significant differences in outcomes on mortality or length of stay (*Table 5*). In addition to that, the number of readmitted patients at the ICU and their severity of illness at readmission based on the SAPS II, and APACHE II and–IV did not significantly differ between the control and the intervention group.

Subgroup-analyses of three age groups (≤49 years, 50–69 years and ≥70 years) or per type of surgery did not result in significant differences between groups.

Table 5. Clinical outcomes of all included patients and patients readmitted to ICU.

Patients overall	**Control group**	**Intervention group**	**p-value**
Unplanned (acute) surgery (%)	57 (18.1)	37 (13.6)	0.16
Length of stay days, median (IQR)			
Hospital	12 (10)	11 (8)	0.39
ICU previous to HDU	1.1 (1.0)	1.1 (1.0)	0.07
HDU	7.5 (7.0)	7.1 (6.0)	0.59
Mortality, number (%)			
In-hospital	5 (1.6)	3 (1.1)	0.9
28-day	7 (2.2)	2 (0.7)	0.27
Readmission ICU, number (%)	43 (13.4)	29 (10.6)	0.36
Readmitted patients	**Control group**	**Intervention group**	**p-value**
Length of stay days, median (IQR)			
Hospital	25 (43)	28 (26)	0.55
ICU previous to HDU	1 (1.1)	1.1 (1.0)	0.14
HDU	3.3 (4.2)	2.4 (4.8)	0.32
ICU readmission	3.4 (9.3)	3.7 (3.6)	0.57
Mortality, number (%)			
In-hospital	4 (9.3)	2 (6.9)	1.0
28-day	5(11.6)	0 (0)	0.15

DISCUSSION

The results of this retrospective study in a surgical high-dependency unit show that the use of an automated MEWS system improves the recording separate MEWS elements and complete MEWS assessments, as well as the resulting adherence to the MEWS hospital protocol. The use of this system improved the MEWS hospital protocol adherence for MEWS assessments using conventional methods during the intervention period. After implementing the automated MEWS system, 39% of the assessments were performed during the intervention period.

Although there was a trend towards improved clinical outcomes in this period, this study did not show significant differences in mortality, length of stay, ICU readmission rate, or severity of illness at ICU readmission. Similar to other studies, our study demonstrated improved accuracy and completeness in the recording of vital signs and complete MEWS after implementation of an automated MEWS system.[13, 14] These positive results on registration outcomes are important since valuable, reliable data become available for research, as quality of care indicators, and as input for clinical decision support system to improve clinical attendance and patient survival.[1, 11, 13-16, 23]

In contrast to similar studies, however, this study did not find a significant improvement of clinical outcomes.[1, 11, 13, 14] The most important difference between this work and others who have found such improvement was the practical and logistical use of the automated MEWS system. The implementation and data management of the automated MEWS in this study was different from previous studies.[1, 13–16, 23] For example, in this study, the automated MEWS system functioned as a CDSS (clinical decision support system) on the bedside and results needed to be copied manually into the hospital EMR system. The lack of direct integration with the EMR and also use of the conventional method during the intervention period were in contrast to previous studies where data collection changed from paper records during the control period to digital recording with automated EWS for the intervention period. In addition, some previous studies applied automated EWS assessments to provide clear overviews with trends on big screens or connections to beepers of ward physicians and RRT.[1, 13–16, 23] These differences in integration of automated EWS systems in daily clinical practice might explain the differences in clinical outcomes.

Second, several previous studies vary in methods used to assess the EWS, like single parameter scoring, EWS, and MEWS.[1, 11, 13, 14] Consequently, this heterogeneity prevents adequate comparison of clinical outcomes.

Third, the present study focused on non-elderly, high-risk surgical patients, while previously studied populations were mainly performed on general surgical and medical wards, and found positive effects in the elderly.[1, 11, 13, 14]

Fourth, the difference in outcomes might be due to the variation in study design. Observation periods between studies differed and could have led to educational or Hawthorne effects in studies with relatively short intervention periods.[11, 14] Additionally, prospective research may be more prone for higher acuity or over-triage of less severe ill patients for ICU admission, possibly leading to diversion of real practice during the intervention periods. For example, the largest prospective multi-center study reporting positive clinical outcomes of Bellomo et al. readmitted more patients to the ICU while the need for ICU interventions like vasopressors, arterial lines and mechanical ventilation in their intervention period was significantly lower compared to the control period.[11] Even though this is explained as a result of earlier recognition of the deteriorating patient, one should be aware that this might also be explained by over-triage. This retrospective study was not able to find significant differences in severity of illness, ICU interventions or clinical outcomes in the overall study population and ICU readmissions. Therefore, retrospective long term designed studies may provide important insights on the effect of automated EWS systems while preventing diversion of real practice.[24]

The retrospective nature of this study in inherent for the limitations of having missing data. Assessments might have been performed without being recorded in the EMR or automated MEWS system. Another limitation is in the size of the groups analysed. If the study were to be powered based on mortality alone, there would have been a need for increasing patient cohort size from approximately 500 to more than 10,000 patients. In addition to that, the administrative burden for the nursing staff can result in missing data due to the automated MEWS system lacking interoperability with the EMR and hospital's computer server. This is likely to be an important reason why only a trend towards improved clinical outcomes was found in this retrospective real-life cohort study.

Automated EWS systems that provide more complete and accurate recording of data have a great potential for future clinical decision support systems and early deterioration detection. Especially if additional data is included such as laboratory, pharmaceutical, and historical data, in combination with the upcoming use of machine learning and artificial intelligence for department or patient-specific MEWS algorithms.[25-29] Therefore, interoperability between automated MEWS systems and the EMR or other medical devices seems essential to prevent solely data input without data management and valuable output to save time for medical staff while achieving more consistent improved clinical outcomes.

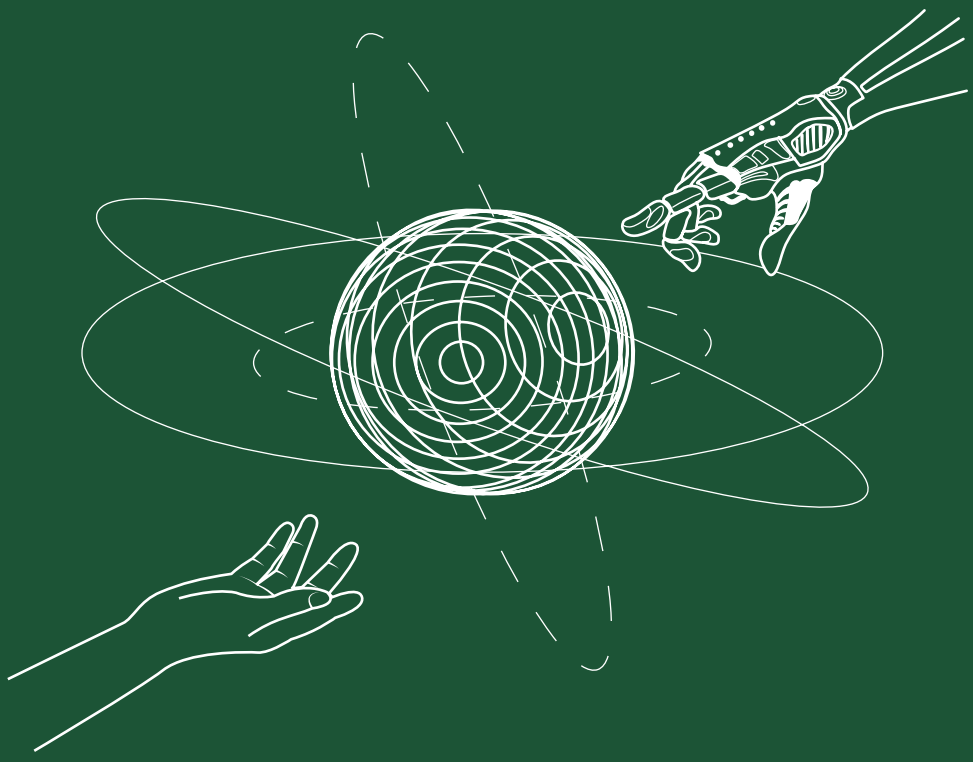
CONCLUSION

Implementation of an automated MEWS on a surgical high dependency unit improves the number of complete MEWS, registered vital signs, and adherence to the local MEWS hospital protocol. However, this positive effect did not translate into a significant decrease in mortality, hospital length of stay, or ICU readmissions. Future research and development on automated EWS systems should focus on data management and technology interoperability to provide actionable insights to the right person at the right time.

REFERENCES

1. Subbe CP, Duller B, Bellomo R. Effect of an automated notification system for deteriorating ward patients on clinical outcomes. *Critical care* 2017; **21**: 52
2. Pearse RM, Moreno RP, Bauer P, et al. Mortality after surgery in Europe: a 7 day cohort study. *Lancet* 2012; **380**: 1059-65
3. Abbott TEF, Ahmad T, Phull MK, et al. The surgical safety checklist and patient outcomes after surgery: a prospective observational cohort study, systematic review and meta-analysis. *British journal of anaesthesia* 2018; **120**: 146-55
4. Alam N, Hobbelink EL, van Tienhoven AJ, van de Ven PM, Jansma EP, Nanayakkara PW. The impact of the use of the Early Warning Score (EWS) on patient outcomes: a systematic review. *Resuscitation* 2014; **85**: 587-94
5. Franklin C, Mathew J. Developing strategies to prevent inhospital cardiac arrest: analyzing responses of physicians and nurses in the hours before the event. *Critical care medicine* 1994; **22**: 244-7
6. De Meester K, Das T, Hellemans K, et al. Impact of a standardized nurse observation protocol including MEWS after Intensive Care Unit discharge. *Resuscitation* 2013; **84**: 184-8
7. Moon A, Cosgrove JF, Lea D, Fairs A, Cressey DM. An eight year audit before and after the introduction of modified early warning score (MEWS) charts, of patients admitted to a tertiary referral intensive care unit after CPR. *Resuscitation* 2011; **82**: 150-4
8. Paterson R, MacLeod DC, Thetford D, et al. Prediction of in-hospital mortality and length of stay using an early warning scoring system: clinical audit. *Clinical medicine* 2006; **6**: 281-4
9. Subbe CP, Williams E, Fligelstone L, Gemmell L. Does earlier detection of critically ill patients on surgical wards lead to better outcomes? *Annals of the Royal College of Surgeons of England* 2005; **87**: 226-32
10. Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified Early Warning Score in medical admissions. *QJM : monthly journal of the Association of Physicians* 2001; **94**: 521-6
11. Bellomo R, Ackerman M, Bailey M, et al. A controlled trial of electronic automated advisory vital signs monitoring in general hospital wards. *Critical care medicine* 2012; **40**: 2349-61
12. Gao H, McDonnell A, Harrison DA, et al. Systematic review and evaluation of physiological track and trigger warning systems for identifying at-risk patients on the ward. *Intensive care medicine* 2007; **33**: 667-79
13. Schmidt PE, Meredith P, Prytherch DR, et al. Impact of introducing an electronic physiological surveillance system on hospital mortality. *BMJ quality & safety* 2015; **24**: 176-7
14. Jones S, Mullally M, Ingleby S, Buist M, Bailey M, Eddleston JM. Bedside electronic capture of clinical observations and automated clinical alerts to improve compliance with an Early Warning Score protocol. *Critical care and resuscitation : journal of the Australasian Academy of Critical Care Medicine* 2011; **13**: 83-8
15. Evans RS, Kuttler KG, Simpson KJ, et al. Automated detection of physiologic deterioration in hospitalized patients. *Journal of the American Medical Informatics Association : JAMIA* 2015; **22**: 350-60

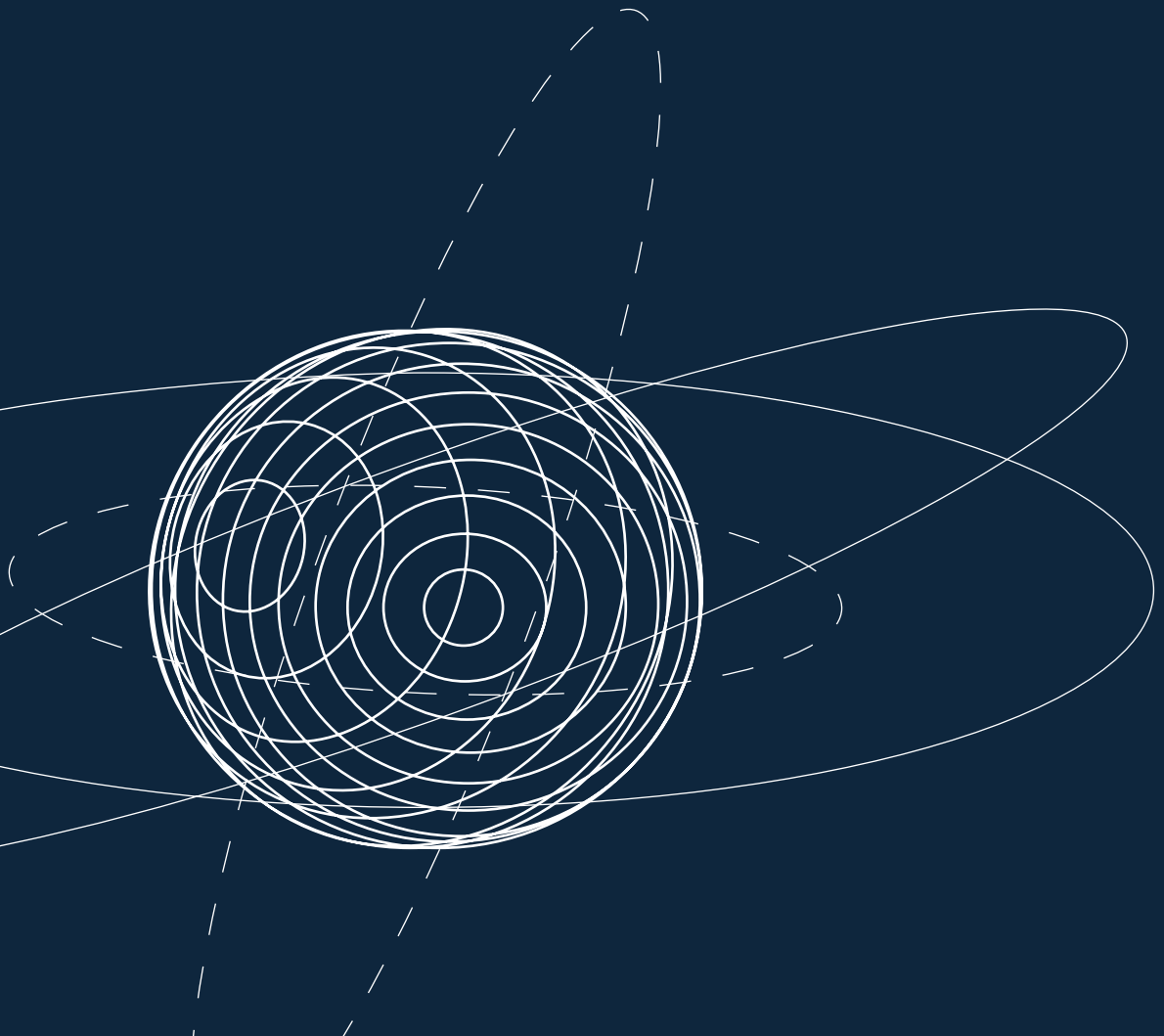
16. Huh JW, Lim CM, Koh Y, et al. Activation of a medical emergency team using an electronic medical recording-based screening system*. *Critical care medicine* 2014; **42**: 801-8
17. Ludikhuize J, Smorenburg SM, de Rooij SE, de Jonge E. Identification of deteriorating patients on general wards; measurement of vital parameters and potential effectiveness of the Modified Early Warning Score. *Journal of critical care* 2012; **27**: 424 e7-13
18. Ludikhuize J, Borgert M, Binnekade J, Subbe C, Dongelmans D, Goossens A. Standardized measurement of the Modified Early Warning Score results in enhanced implementation of a Rapid Response System: a quasi-experimental study. *Resuscitation* 2014; **85**: 676-82
19. Dawes TR, Cheek E, Bewick V, et al. Introduction of an electronic physiological early warning system: effects on mortality and length of stay. *British journal of anaesthesia* 2014; **113**: 603-9
20. McGain F, Cretikos MA, Jones D, et al. Documentation of clinical review and vital signs after major surgery. *The Medical journal of Australia* 2008; **189**: 380-3
21. Hands C, Reid E, Meredith P, et al. Patterns in the recording of vital signs and early warning scores: compliance with a clinical escalation protocol. *BMJ quality & safety* 2013; **22**: 719-26
22. Corp I. IBM SPSS Statistics for Windows; Version 23.0. Armonk, NY. 2015
23. Kollef MH, Chen Y, Heard K, et al. A randomized trial of real-time automated clinical deterioration alerts sent to a rapid response team. *Journal of hospital medicine* 2014; **9**: 424-9
24. Cohen AT, Goto S, Schreiber K, Torp-Pedersen C. Why do we need observational studies of everyday patients in the real-life setting? *European Heart Journal Supplements* 2015; **17**: D2-D8
25. Florath I, Albert A, Hassanein W, et al. Current determinants of 30-day and 3-month mortality in over 2000 aortic valve replacements: Impact of routine laboratory parameters. *European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery* 2006; **30**: 716-21
26. Loekito E, Bailey J, Bellomo R, et al. Common laboratory tests predict imminent death in ward patients. *Resuscitation* 2013; **84**: 280-5
27. Khurana HS, Groves RH, Jr., Simons MP, et al. Real-Time Automated Sampling of Electronic Medical Records Predicts Hospital Mortality. *The American journal of medicine* 2016; **129**: 688-98 e2
28. Hackmann G, Chen M, Chipara O, et al. Toward a two-tier clinical warning system for hospitalized patients. *AMIA Annual Symposium proceedings AMIA Symposium* 2011; **2011**: 511-9
29. Petit C, Bezemer R, Atallah L. A review of recent advances in data analytics for post-operative patient deterioration detection. *Journal of clinical monitoring and computing* 2018; **32**: 391-402



PART 4

A CDS system based on a complex transparent model with an automated, closed-loop, functionality

4.1



Intelligent dynamic clinical checklists improved checklist compliance in the intensive care unit

A.J.R. De Bie, S. Nan, L.R.E. Vermeulen, P.M.E. Van Gorp, R.A. Bouwman, A.J.G.H. Bindels, and H.H.M. Korsten

British journal of anaesthesia 2017

ABSTRACT

Background

Checklists can reduce medical errors. However, the effectiveness of checklists is hampered by lack of acceptance and compliance. Recently, a new type of checklist with dynamic properties has been created to provide more specific checklist items for each individual patient. Our purpose in this simulation-based study was to investigate a newly developed intelligent dynamic clinical checklist (DCC) for the intensive care unit (ICU) ward round.

Methods

Eligible clinicians were invited to participate as volunteers. Highest achievable scores were established for six typical ICU scenarios to determine which items must be checked. The participants compared the DCC with the local standard of care. The primary outcomes were the caregiver satisfaction score and the percentages of checked items overall and of critical items requiring a direct intervention.

Results

In total, 20 participants were included, who performed 116 scenarios. The median percentage of checked items was 100.0% with the DCC and 73.6% for the scenarios completed with local standard of care ($P<0.001$). Critical items remained unchecked in 23.1% of the scenarios performed with local standard of care and 0.0% of the scenarios where the DCC was available ($P<0.001$). The mean satisfaction score of the DCC was 4.13 out of 5.

Conclusion

This simulation study indicates that an intelligent DCC significantly increases compliance with best practice by reducing the percentage of unchecked items during ICU ward rounds, while the user satisfaction rate remains high. Real-life clinical research is required to evaluate this new type of checklist further.

INTRODUCTION

In America, it has been estimated that the deaths of 210 000 hospitalized patients are associated with preventable adverse events each year.[1] This large number can be explained if one considers that most medical procedures are still based on human memory.[2, 3] To prevent these adverse events, a huge diversity of medical guidelines and protocols have been introduced, but it remains a challenge to implement them in daily practice. For example, only 56% of patients in the intensive care unit (ICU) are treated according to the best practice for which they are eligible.[4] To overcome these problems, a benchmark used in other high-risk industries, the checklist, has been tested as a method in medical care, with encouraging results.[5–10] Haynes and colleagues[5] showed that the surgical safety checklist standardizes preoperative care, resulting in a cost-effective reduction of morbidity and mortality. Likewise, De Vries and colleagues[7] demonstrated that implementing multidisciplinary checklists in the surgical pathway, from admission to discharge, significantly reduced the proportion of patients with one or more complications from 15.4 to 10.6% in Dutch hospitals.

However, numerous subsequent qualitative studies could not reproduce these beneficial effects, which could be attributable to the remaining challenge of checklist implementation in medical care, which is a lack of acceptance and compliance.[5, 7, 11–16] A possible cause could be that current static checklists negatively interfere with the daily workflow of caregivers because they do not provide contextual information that makes it easier to complete the checklist and they cannot include or exclude items based on the characteristics of a particular patient and caregiver.

Recently, Nan and colleagues[17] created TraceBook, a new decision support system that integrates workflow management with the use of dynamic clinical checklists (DCCs) in a process oriented and context-aware manner to make clinical processes more traceable and the people in it more accountable. These new forms of intelligent checklists derive their dynamic property from being connected with the electronic health record (EHR) and other electronic medical databases. These checklists are therefore able to provide real-time relevant information and specific items of patients to the specific user. Our hypothesis is that these dynamic characteristics can ensure a high satisfaction rate among clinicians and improve the compliance with best eligible practice.

The aim of this study was to evaluate whether the compliance with best eligible practice is increased with this new type of checklist, while keeping the satisfaction rate high.

METHODS

This simulation-based study was conducted in November 2014 in the Intensive Care Department of Catharina Hospital Eindhoven, a tertiary hospital in The Netherlands. The simulations were performed as in situ simulations in a real room of the ICU with a mannequin as the patient.

Scenario development

We created six patient scenarios based on data of patients who had been admitted to the ICU and deliberately implemented some flaws (Supplementary Appendix 1). The patients were virtually admitted in the EHR-test environment (CS-EZIS test, Chipsoft BV, Amsterdam, The Netherlands).

For each scenario, we established a highest achievable score containing all the items that should be checked by the participant during each ward round. The items were identified based on guidelines, the current paper checklist (Supplementary Appendix 2) and local expert opinion. Medical issues requiring a direct intervention were called critical items. The scenarios with their corresponding highest achievable scores were reviewed and approved by two intensivists (A.J.G.H.B. and H.H.M.K.) of the research team, who did not participate in the trial.

Study participants

Clinically active clinicians were eligible to participate if they had ward round experience on the ICU for at least 1 month between January 2013 and November 2014. Participants could be intensivists, nurse practitioners of the ICU, residents, or final year medical students after an ICU internship. Eligible participants were invited to participate, and participation was voluntary. When completing the survey, participants gave verbal and written consent for the use of the collected data for publication.

Local standard of care

The current local standard of care (LSC) during an ICU ward round is a paper checklist that is available at the bedside to be used at the caregiver's convenience. This paper checklist is based on the FAST HUG mnemonic,[3, 18] and since its introduction on the ICU, intensivists have optimized this checklist by adding extra items (Supplementary Appendix 2).

For more than a decade, the Catharina Hospital Eindhoven has also been using the clinical decision support system (CDSS) GASTON to improve guideline compliance

regarding medication.[19–21] This CDSS is connected to the EHR and checks predetermined pharmacological clinical rules for the ICU (Supplementary Appendix 3). If these clinical rules are violated, the CDSS produces alerts.[20] An example of such a violation could be a patient on the ICU receiving non-steroidal anti-inflammatory drugs without gastric protection. Once a day, after the ICU ward rounds, a list of all the alerts is generated and evaluated by a hospital pharmacist, who then contacts the physician on duty by telephone to discuss the recommendations. This physician decides whether a recommendation should lead to an intervention or not.[20]

Intelligent dynamic clinical checklist

The intervention was based on the use of an intelligent DCC that generates a dedicated checklist for each individual patient. To do this, the systems of TraceBook and GASTON both use a rule engine containing a model of algorithms, comparable with a decision tree, with general clinical rules and pharmacological rules that are both specifically applicable to the ICU.[19, 20] First GASTON gathers the relevant information about the patient from different medical information systems, such as patient monitors, the EHR, the pharmaceutical prescription system, and others. Then GASTON and TraceBook run the rule engines containing the clinical and pharmaceutical rules with their algorithms, and TraceBook determines which rules are relevant for a specific patient in a specific context and should become a checkable item for the DCC of that particular patient. Some of these items can be checked automatically, depending on the available information, on the algorithm of the rules, and on whether local consensus of the professionals decided that a rule may be checked automatically. This last condition also implies that professionals can decide that some rules should not be checked automatically.

The model for the DCC for the ICU ward round is based on the combination of our local paper checklist, which is also available during LSC, and the pharmacological rules that are specifically applicable for our ICU and generated by GASTON (Supplementary Appendixes S2 and S3).

Figure 1 provides a schematic overview of how a DCC is composed, showing a small part of the algorithm for prescribing analgesia based on the pain rating scale, because this comprehensively illustrates how the clinical rules work and how they generate checkable or automatically checkable items in the DCC. *Figure 1* also demonstrates a part of the DCC where TraceBook can highlight text for extra attention and provide the user with data from the EHR and guidelines on request.

The whole system was designed to create or modify the rules in the model easily. No rules were adjusted, added, or removed during the simulation procedure. The number

of items and critical items that were relevant and needed to be checked per scenario are described in *Table 1*. In addition, *Table 1* shows the number of these relevant checkable items that can be checked automatically by the DCC.

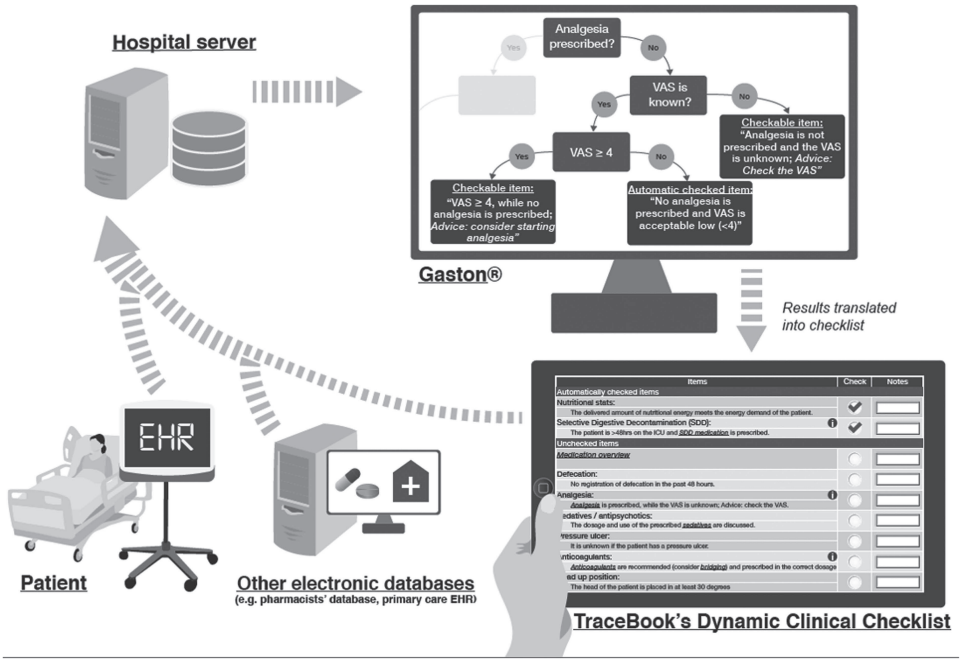


Figure 1. A schematic overview of how a TraceBooks DCC is composed.
The information about a patient is gathered by GASTON on the hospital server. The rule engines of GASTON and TraceBook decide, based on algorithms, which rule is relevant and can become a checkable item or automatically checked item. A small part of the algorithm of prescribing analgesia based on the pain rating scale is shown, which can provide an automatically checked item if no analgesia is prescribed and the pain rating score is low (<4). DCC, dynamic clinical checklist; e.g., example; EHR, electronic health record; VAS, visual analogue scale.

Simulation procedure

Participants were randomly assigned into two groups for a crossover design. Group 1 performed Scenarios 1–3 by local standard of care, followed by a tutorial about the DCC, and then they completed Scenarios 4–6 with the DCC available. Group 2 performed Scenarios 4–6 by local standard of care, followed by the same tutorial, and then they accomplishing Scenarios 1–3 with the DCC available (*Figure 2*).

Table 1. Characteristics of the simulation scenarios.
The table includes the patient characteristics, the number of checkable items overall, critical items, and items that can be checked automatically; and percentages of these automatically checkable items that were checked in the scenarios performed by local standard of care or with the DCC available.

Scenario	**1**	**2**	**3**	**4**	**5**	**6**
Gender	Female	Male	Male	Male	Female	Female
Age (years)	62	66	61	74	68	42
APACHE-II at admission	19	26	18	11	32	26
Mechanical ventilation	Yes	Yes	Yes	No	No	Yes
Continuous sedation (RASS)	Yes (3)	Yes (-5)	Yes (-4)	No	No	Yes (-5)
Central venous line in situ	Yes	Yes	Yes	Yes	Yes	Yes
Items						
Checkable items (n)	31	30	29	23	29	27
Checkable critical items (n)	14	11	9	5	13	11
Automatically checkable items (n)	3	3	4	3	4	2
Percentage of the automatically checkable items that were checked in scenarios completed by local standard of care (%)	82	94	85	52	93	50
Percentage of the automatically checkable items that were checked in scenarios completed with the DCC (%)	100	100	100	100	100	100

APACHE, Acute Physiology and Chronic Health Evaluation; DCC, dynamic clinical checklist; LSC, local standard of care; RASS, Richmond Agitation–Sedation Scale

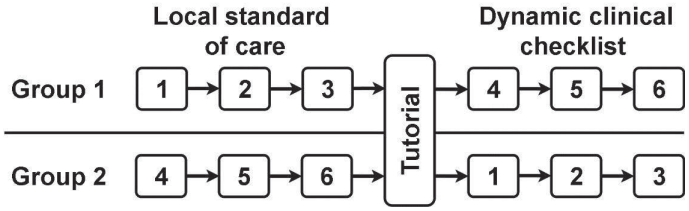


Figure 2. Crossover design of the simulation study, with the order of scenarios performed by the two groups.

As in daily routine, the principal investigator informed each participant about the clinical history of each simulated scenario, including medical history, physical examination, diagnostic tests, and the conclusion with the plan for the day. After this presentation, the participant had the opportunity to agree with the proposed plan or to adjust it as

he preferred. To make this decision, the participant could choose to use either the paper checklist or the DCC, depending on which one was available in the scenario, or not to use a checklist. The scenario was considered complete when the participant declared that he had finished the scenario. After finishing all six scenarios, the participant completed a survey containing questions on usability, training and support, behaviour change, usefulness, and user satisfaction on a five-point Likert scale (with 1 totally disagree, 2 disagree, 3 neutral, 4 agree, and 5 totally agree). Participants were also asked to rate their satisfaction of the DCC on a scale from 1 to 5, where a higher score indicates better satisfaction (Supplementary Appendix 4).

Data collection and analyses

All scenarios were observed by one observer and recorded on video. The observer was sitting out of sight of the participants and noted which items were checked. Items could be checked verbally or in writing, and interventions were documented. The principal investigator reviewed all video recordings to doublecheck which items had been checked.

The primary outcomes were the satisfaction rate of the DCC and the percentages of checked items and unchecked critical items during the scenarios. The secondary outcomes were the required time from the end of the presentation until the end of the scenario and the percentage of scenarios needing a telephone call by the pharmacist based on violated pharmacological clinical rules.

Statistical analyses were performed with SPSS version 21 (IBM Corp., Armonk, NY, USA). The distribution of continuous variables was assessed with Kolmogorov–Smirnov tests. The χ^2-test and independent-samples t-test were used if data were parametric, whereas the Mann–Whitney U-test was used for non-parametric data. A two-sided P-value <0.05 was considered statistically significant.

RESULTS

Participants and scenarios

Twenty clinicians consented to participate in this study: three intensivists, 15 residents, one nurse practitioner, and one final year medical student. The difference in experience (in weeks) between Group 1 [median=20, interquartile range (IQR) 16 – 52] and Group 2 (median = 54, IQR 16 – 200) was not significant ($p = 0.23$). In total, the participants completed 116 scenarios. Two participants could not fulfil all six scenarios because of work-related issues and performed four scenarios instead. In one instance, the DCC had

been forgotten, and therefore this simulated scenario was counted as a ward round performed with the local standard of care. The patient characteristics of each scenario are described in *Table 1*.

Outcomes

Figure 3 illustrates the comparison between the scenarios performed with LSC and the scenarios accomplished with the DCC, showing an increase of the median percentage of checked items from 73.6% (IQR 64.5 – 79.3) to 100% (IQR 100.0 – 100.0; p<0.001, with z = -7.74). The median percentage of unchecked critical items decreased from 23.1% (IQR 9.0 – 40.0) to 0.0% (IQR 0.0 – 0.0; p<0.001, with z = 9.61). *Table 1* describes the percentages of automatically checkable items that were checked per scenario if LSC or the DCC was applied.

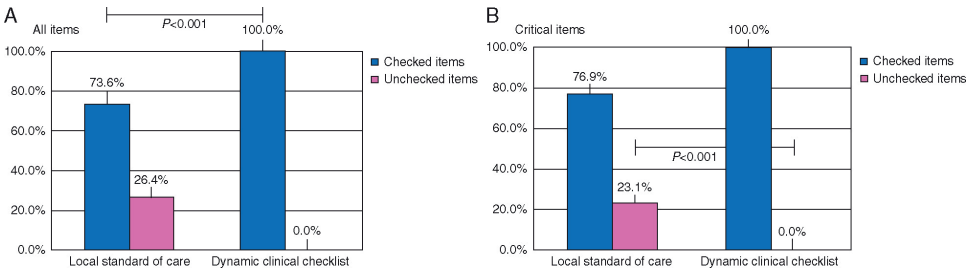


Figure 3. Median percentages of checked items overall (A) and unchecked critical items (B), needing a direct intervention, during simulated intensive care ward rounds with only a paper checklist available or also a dynamic clinical checklist available.

Based on CDSS alerts after the ward round, the pharmacist had to call after 80.0% of the scenarios performed with LSC, compared with 3.6% (p<0.001) of the scenarios performed with the DCC available (*Figure 4*).

For four scenarios, the time from the end of the presentation until the end of the scenario was shorter with LSC than with the DCC [264 (SD 135) vs 364 (125) seconds ; p<0.001, 95% confidence interval, -150 to -51]. In two of the scenarios, no significant difference in time was perceived (*Figure 5*).

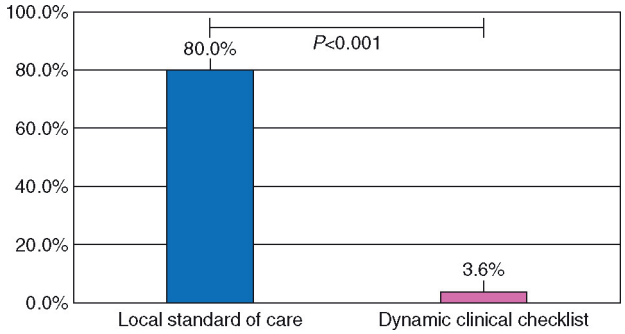


Figure 4. Median percentage of scenarios requiring a pharmacist's call after the scenario, owing to violated pharmaceutical clinical rules, in the scenarios with only a paper checklist available and the group with a dynamic clinical checklist available.

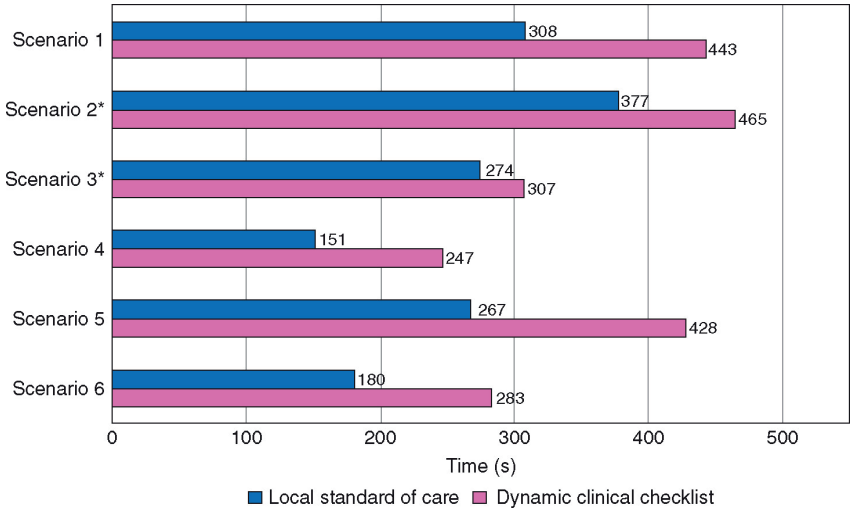


Figure 5. The median time (in seconds) between the end of scenario presentation until the end of the scenario for each of the six scenarios.

There was no significant difference in time in two scenarios (*)

The mean satisfaction score of the DCC was 4.13 out of 5 (95% confidence interval of 3.91 – 4.34). All participants agreed with the statement that there is a potential for intelligent DCCs in medical care. These last two results are described with the other results of the survey in Supplementary Appendix 4.

DISCUSSION

In this prospective simulation-based study, we observed that the compliance with the best eligible practice during ICU ward rounds improved if an intelligent DCC was available, based on a significantly improved percentage of checked items and a significantly reduced percentage of unchecked critical items. This improvement significantly reduced the need for intervention recommendations by the hospital pharmacist after the ward rounds. Although the time required to complete the scenarios with the DCC was significantly longer in four of the six scenarios, the satisfaction score for the DCC was high.

The most notable outcome of our study is that with the DCC, the median percentage of checked items was 100%, as opposed to 73.6% with LSC. The latter percentage is similar to the percentage of checked items found in other studies that used paper checklists.[8 22–24] Our results with the DCC cannot be compared with the results of other studies, because the intelligent DCC is a new sophisticated form of checklist. Therefore, research on this particular type of checklist is not available, and research on digital checklists overall is scarce.

Thongprayoon and colleagues[22] showed that if a digital checklist was used during ICU ward rounds instead of a paper checklist with identical questions, the percentage of unchecked items decreased from 14.9 to 8.8%. In our study, an even larger reduction was established. This can probably be explained by the dynamic design of the DCC, with features such as items being checked automatically and providing valuable information so that the checklist can be completed more easily. However, the comparison between these two studies should be considered with care, as the checklists used in the two studies also differ in terms of content.

Our observation of a significantly longer time needed to complete ward rounds with digital checklists is consistent with the results of other studies, with only one study finding no difference of time.[22, 25, 26] However, the extra time required was never >3 min. Besides, this longer duration can be explained by the increased number of detected errors that were resolved. In the long run, this will probably prevent complications and errors, which commonly require more time of caregivers. This hypothesis is supported by the fact that the use of the DCC significantly reduced the number of CDSS alerts, which would have required the hospital pharmacist to recommend interventions after the ward rounds.

The detected high satisfaction score of the DCC is supported by studies reporting an improved checklist usefulness, workload, and integration in workflow when a digital checklist was used instead of a paper checklist. However, evidence on differences in user satisfaction rates between both forms of checklists is lacking.[22, 27] Based on the results of the present study, we think that the intelligent DCC can achieve a high satisfaction rate among caregivers and could therefore challenge the practical downsides of current static checklists that may be responsible for low checklist compliance. This is important because there seems to be a direct relationship between checklist compliance and morbidity reduction.[15, 28] A likely explanation for the high satisfaction rate of the DCC could be the direct experience of benefit for the user, because the DCC acts as a cognitive aid and helps the user to complete the checklist. This ensures that the DCC becomes a helpful tool for clinicians, instead of being a mandatory, workload increasing tool that has beneficial effects only outside of the user's scope. Another advantage of a DCC generated with the TraceBook system is that the clinical rules can easily be updated or modified, which answers the concern that current static checklists are too slow to adapt to improvements in medical practice.[28]

The most important limitation of our study is inherent to the simulation-based study design. Although the testing environment was a room of the ICU with a mannequin and EHR available, common distractions on an ICU were missing, with no real-life patient, nursing staff, or family available for the participant to gather information from. A mannequin was used because a constant performance as a realistic intensive care patient for a more expensive actor is difficult and could introduce too much variation in performance or distract the participants from the interventions that needed to be investigated.

Another limitation of our study is that all scenarios were new for the participants, whereas normally the physicians are more or less aware of the patients' conditions before starting their ward rounds. Moreover, all the scenarios were presented, as objectively as possible, by the same principal investigator, who was involved in the development of the DCC. This may have had impacts on the participants' performances that have not been evaluated during our study, and it is possible that participants tried to please the investigator while completing the surveys. Nonetheless, in highly reliable organizations that use checklists, simulation is indispensable for testing and revising checklists.[2] Simulation is therefore also accepted in medicine as a method for evaluating the effectiveness of new clinical tools.[2, 29–31] The DCC also is a new computer-based tool, and the impact of these features on the results of the present study remains unclear. As two final limitations, we evaluated the compliance during one ICU ward round of one patient rather than several, and we assessed the satisfaction score of the DCC, but

not of the LSC. Therefore, our results shed no light on the long-term compliance and satisfaction with the DCC, nor on the comparison between the satisfaction scores of the DCC and of paper checklists.

More research is needed in a real-world clinical setting over a longer period of time to investigate the long-term compliance and satisfaction rate of the DCC. In addition, it would be interesting to evaluate how the use of different DCCs by different types of medical staff in clinical pathways can improve the traceability of medical processes, the accountability of medical staff, and the safety of medical care.

CONCLUSION

Our simulation-based study indicates that using an intelligent DCC during ICU ward rounds improves compliance with best eligible practice based on a reduction of unchecked critical items, while user satisfaction ratings are high. Therefore, the intelligent DCC has the potential to become a helpful tool for clinicians while improving patient safety. More research is needed to evaluate this new type of intelligent checklist in real clinical settings over longer periods of time.

REFERENCES

1. James JT. A new, evidence-based estimate of patient harms associated with hospital care. *J Patient Saf* 2013; **9**: 122–8
2. Thomassen Ø, Espeland A, Søfteland E, Lossius HM, Heltne JK, Brattebø G. Implementation of checklists in health care; learning from high-reliability organisations. *Scand J Trauma Resuscitation Emerg Med* 2011; **19**: 53
3. Vincent JL. Give your patient a fast hug (at least) once a day. *Crit Care Med* 2005; **33**: 1225–9
4. Ilan R, Fowler RA, Geerts R, Pinto R, Sibbald WJ, Martin CM. Knowledge translation in critical care: factors associated with prescription of commonly recommended best practices for critically ill patients. *Crit Care Med* 2007; **35**: 1696–702
5. Haynes AB, Weiser TG, Berry WR, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2009; **360**: 491–9
6. Blot K, Bergs J, Vogelaers D, Blot S, Vandijck D. Prevention of central line-associated bloodstream infections through quality improvement interventions: a systematic review and meta-analysis. *Clin Infect Dis* 2014; **59**: 96–105
7. de Vries EN, Prins HA, Crolla RM, et al. Effect of a comprehensive surgical safety system on patient outcomes. *N Engl J Med* 2010; **363**: 1928–37
8. Conroy KM, Elliott D, Burrell AR. Testing the implementation of an electronic process-of-care checklist for use during morning medical rounds in a tertiary intensive care unit: a prospective before–after study. *Ann Intensive Care* 2015; **5**: 60
9. Brunsveld-Reinders AH, Arbous MS, Kuiper SG, de Jonge E. A comprehensive method to develop a checklist to increase safety of intra-hospital transport of critically ill patients. *Crit Care* 2015; **19**: 214
10. Nama A, Sviri S, Abutbul A, Stav I, van Heerden PV. Successful introduction of a daily checklist to enhance compliance with accepted standards of care in the medical intensive care unit. *Anaesth Intensive Care* 2016; **44**: 498–500
11. de Vries EN, Hollmann MW, Smorenburg SM, Gouma DJ, Boermeester MA. Development and validation of the SURgical PAtient Safety System (SURPASS) checklist. *Qual Saf Health Care* 2009; **18**: 121–6
12. Fourcade A, Blache JL, Grenier C, Bourgain JL, Minvielle E. Barriers to staff adoption of a surgical safety checklist. *BMJ Qual Saf* 2012; **21**: 191–7
13. Borchard A, Schwappach DL, Barbir A, Bezzola P. A systematic review of the effectiveness, compliance, and critical factors for implementation of safety checklists in surgery. *Ann Surg* 2012; **256**: 925–33
14. Urbach DR, Govindarajan A, Saskin R, Wilton AS, Baxter NN. Introduction of surgical safety checklists in Ontario, Canada. *N Engl J Med* 2014; **370**: 1029–38
15. van Klei WA, Hoff RG, van Aarnhem EE, et al. Effects of the introduction of the WHO "Surgical Safety Checklist" on in-hospital mortality: a cohort study. *Ann Surg* 2012; **255**: 44–9

16. Cavalcanti AB, Bozza FA, Machado FR, et al. Effect of a quality improvement intervention with daily round checklists, goal setting, and clinician prompting on mortality of critically ill patients: a randomized clinical trial. *JAMA* 2016; **315**: 1480–90
17. Nan S, Van Gorp PME, Korsten EHM, et al. Tracebook: a dynamic checklist support system. *Comput Base Med* Syst 2014; 48–51. Available from http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=6881324
18. Papadimos TJ, Hensley SJ, Duggan JM, et al. Implementation of the "FASTHUG" concept decreases the incidence of ventilator-associated pneumonia in a surgical intensive care unit. *Patient Saf Surg* 2008; **2**: 3
19. de Clercq PA, Hasman A, Blom JA, Korsten HH. Design and implementation of a framework to support the development of clinical guidelines. *Int J Med Informat* 2001; **64**: 285–318
20. Scheepers-Hoeks AMJWGR, Neef C, Doppen AMJ, Korsten EHM. Preventive prescribing of laxatives for opioid-induced constipation using electronic clinical rule implementation by clinical pharmacists. *Adv Pharmacoepidemiol Drug Saf* 2014; **3**: 1–5
21. van der Linden CM, Jansen PA, van Marum RJ, Grouls RJ, Egberts TC, Korsten EH. An electronic system to document reasons for medication discontinuation and to flag unwanted represcriptions in geriatric patients. *Drugs Aging* 2012; **29**: 957–62
22. Thongprayoon C, Harrison AM, O'Horo JC, Berrios RA, Pickering BW, Herasevich V. The effect of an electronic checklist on critical care provider workload, errors, and performance. J *Intensive Care Med* 2016; **31**: 205–12
23. Pucher PH, Aggarwal R, Almond MH, Darzi A. Surgical care checklists to optimize patient care following postoperative complications. *Am J Surg* 2015; **210**: 517–25
24. de Vries EN, Prins HA, Bennink MC, et al. Nature and timing of incidents intercepted by the SURPASS checklist in surgical patients. *BMJ Qual Saf* 2012; **21**: 503–8
25. Rouse SRW. Computer-based manuals for procedural information. *IEEE Trans Syst Man Cybern Syst* 1980; **10**: 506–10
26. Blike G, Biddle C. Preanesthesia detection of equipment faults by anesthesia providers at an academic hospital: comparison of standard practice and a new electronic checklist. *AANA J* 2000; **68**: 497–505
27. Garg T, Lee JY, Evans KH, Chen J, Shieh L. Development and evaluation of an electronic health record-based best-practice discharge checklist for hospital patients. *Jt Comm J Qual Patient Saf* 2015; **41**: 126–31
28. Tung A. Checklists in critical care: are they needed? *Curr Concepts Adult Crit Care* 2015; **2015**: 117–22
29. Ziewacz JE, Arriaga AF, Bader AM, et al. Crisis checklists for the operating room: development and pilot testing. *J Am Coll Surg* 2011; **213**: 212–7.e10
30. Arriaga AF, Bader AM, Wong JM, et al. Simulation-based trial of surgical-crisis checklists. *N Engl J Med* 2013; **368**: 246–53
31. Watterson L, Flanagan B, Donovan B, Robinson B. Anaesthetic simulators: training for the broader health-care profession. *Aust N Z J Surg* 2000; **70**: 735–7

SUPPLEMENTAL DIGITAL CONTENT 1

Appendix 1. Short description of the six scenarios

Scenario	Day of ICU admission	Short description
1	1	Pneumonia and sepsis
2	3	Out of hospital cardiac arrest, after therapeutic hypothermia
3	2	CABG and MVA, complicated by cardiac stunning
4	1	CABG and atrial valve replacement, uncomplicated
5	3	High output stoma with hypokalaemia after chemotherapy
6	1	Subarachnoid bleeding, complicated by status epilepticus

Abbreviations: ICU = Intensive Care Unit; CABG = coronary artery bypass graft surgery; MVA = Mitral valve annuloplasty.

Appendix 2. Paper checklist

Appendix 2. The paper checklist.

Add Patient sticker

Department IC – Checklist IC rounds

Date: _____/_____/_____

F Enteral feeding possible?
Calories sufficient?
Recent defecation?
A Suitable pain relief?
Can / Should dose be adjusted? (VAS)
S Sedatives and / or antipsychotics prescribed?
Adjust dose? (RASS or CAM-ICU)
D Pressure Ulcer present? Prophylaxis / treatment needed?
T Indication for therapeutic anticoagulant?
Reason for bridging?
Adequate thrombosis prophylaxis?
H Headboard is at least 30 degrees up?
Indication for protective ventilation?
U Ulcer prophylaxis?
G Glucose protocol?
S SDD protocol? (selective intestinal decontamination)

Lines
Place?
Since?
Antibiotics
Since?
Reason? (Cultivate?)
Levels?
Stop date?
Prior history
Does the prior history give reasons to deviate from the normal procedures?
Physical examination
Are there new aspects leading from the physical examination that require policy change?
Laboratory examination
Are there laboratory results that require a change of policy?
Radiological examination
Requires the radiological examination results a change of policy or intervention (position tube, pneumatic, etc.)?
Conclusion
Working diagnosis?
Goal formulated?
Communicated with all involved parties (nurses, consultants, family)?
Complications
Scored in EHR?
NICE data completed?

Appendix 3. All pharmaceutical clinical rules that are checked by CDSS Gaston and the DCC for all scenarios [19-21]

	Clinical rule
1.	The system checks if Methotrexate, with folic acid is administered; If so, it will provide a checkable item to check if dosage is correct and if folic acid is administered.
2.	The system checks if nefrotoxic medication is administered in case of kidney dysfunction; If so, it will provide a checkable item to check if nefrotoxic medication is needed or if the dosage can be changed.
3.	The system checks if laxatives are started simultaneously with the administered opiates; If not, it will provide a checkable item to start laxatives when there are no contraindications
4.	The system checks if aminoglycosides are administered; If so, it will provide a checkable item to check if aminoglycosides levels are monitored and if dosage is correct.
5.	The system checks if there is a hyper- or hypokalium and if so it checks if there is any medication responsible for it. If so, it will provide a checkable item to check potassium levels and medication.
6.	The system checks if there is a hyper- or hyponatrium and if so it checks if there is any medication responsible for it. If so, it will provide a checkable item to check natrium levels and medication.
7.	The system checks if there is a hyper- or hypocalcemia and if so it checks if there is any medication responsible for it. If so, it will provide a checkable item to check calcium levels and medication.
8.	The system checks if stress ulcer prophylaxis is started and checks if NSAIDs are started. If not so, it will provide a checkable item to start stress ulcer prophylaxis and to check if NSAID is necessary.
9.	The system checks if the patient with heart failure gets medication that is contraindicated in heart failure. If so, it will provide a checkable item to check if this medication is necessary and to evaluate if it can be stopped.
10.	The system checks if the INR is >6. If so, it will it will provide a checkable item to suggest to start Vitamin K.
11.	The system checks if Lithium is prescribed for the patient and if blood levels of Lithium are known and acceptable. If so, it will provide a checkable item to suggest checking the blood levels of lithium or if the dosage of Lithium needs to be modified.
12.	The system checks if Digoxin is prescribed for the patient and if blood levels of Digoxin are known and acceptable. If so, it will provide a checkable item to suggest checking the blood levels of Digoxin or if the dosage of Digoxin needs to be modified.
13.	The system checks if Clozapine is prescribed for the patient and if blood levels of Clozapine are known and acceptable. If so, it will provide a checkable item to suggest checking the blood levels of Clozapine or if the dosage of Clozapine needs to be modified.

	Clinical rule
14.	The system checks if Phenytoin is prescribed for the patient and if blood levels of Phenytoin are known and acceptable. If so, it will provide a checkable item to suggest checking the blood levels of Phenytoin or if the dosage of Phenytoin needs to be modified.
15.	The system checks if enteral feeding and levothyroxine are given at the same time. If so, it will provide a checkable item to suggest skip one bolus of enteral feeding or pause enteral feeding for half an hour if given continuously.
16.	The system checks if dalteparin dosage >5000IE/day if the patient is >80kg. If not so, it will provide a checkable item to start daltaparin 5000IE/day.
17.	The system checks if the patient gets Daltaparin and whether the INR is two consecutive times > 2.2 If so, it will provide a checkable item to suggest pausing the Dalteparin.
18.	The system checks if the patient gets Amiodaron 1200mg/24hr >3 days If so, it will provide a checkable item to suggest to correct the dosage to 600mg/24hr or start oral Amiodaron.
19.	The system checks if Vancomycin is prescribed for the patient and if blood levels of Vancomycin are known and acceptable. If so, it will provide a checkable item to suggest checking the blood levels of Vancomycin or if the dosage of Vancomycin needs to be modified.
20.	The system checks if Amikacin is prescribed for the patient and if blood levels of Amikacin are known and acceptable. If so, it will provide a checkable item to suggest checking the blood levels of Amikacin or if the dosage of Amikacin needs to be modified.
21.	The system checks if selective oral decontamination is prescribed for patient admitted on the IC >48 hours. If not so, it will provide a checkable item to suggest starting selective oral decontamination.
22.	The system checks if the patient has an enteral tube and if the prescribed medication is eligible to be given through the enteral tube. If not so, it will provide a checkable item to suggest to change the ineligible medication to medication that can be given intravenously of with the enteral tube.
23.	The system checks if a venous or arterial line is in situ >7 days. If so, it will provide a checkable item to consider change the line or evaluate if the line is still needed.

Appendix 4. Survey that was completed by the participants immediately after all the ward rounds with the results (in percentages and absolute value)

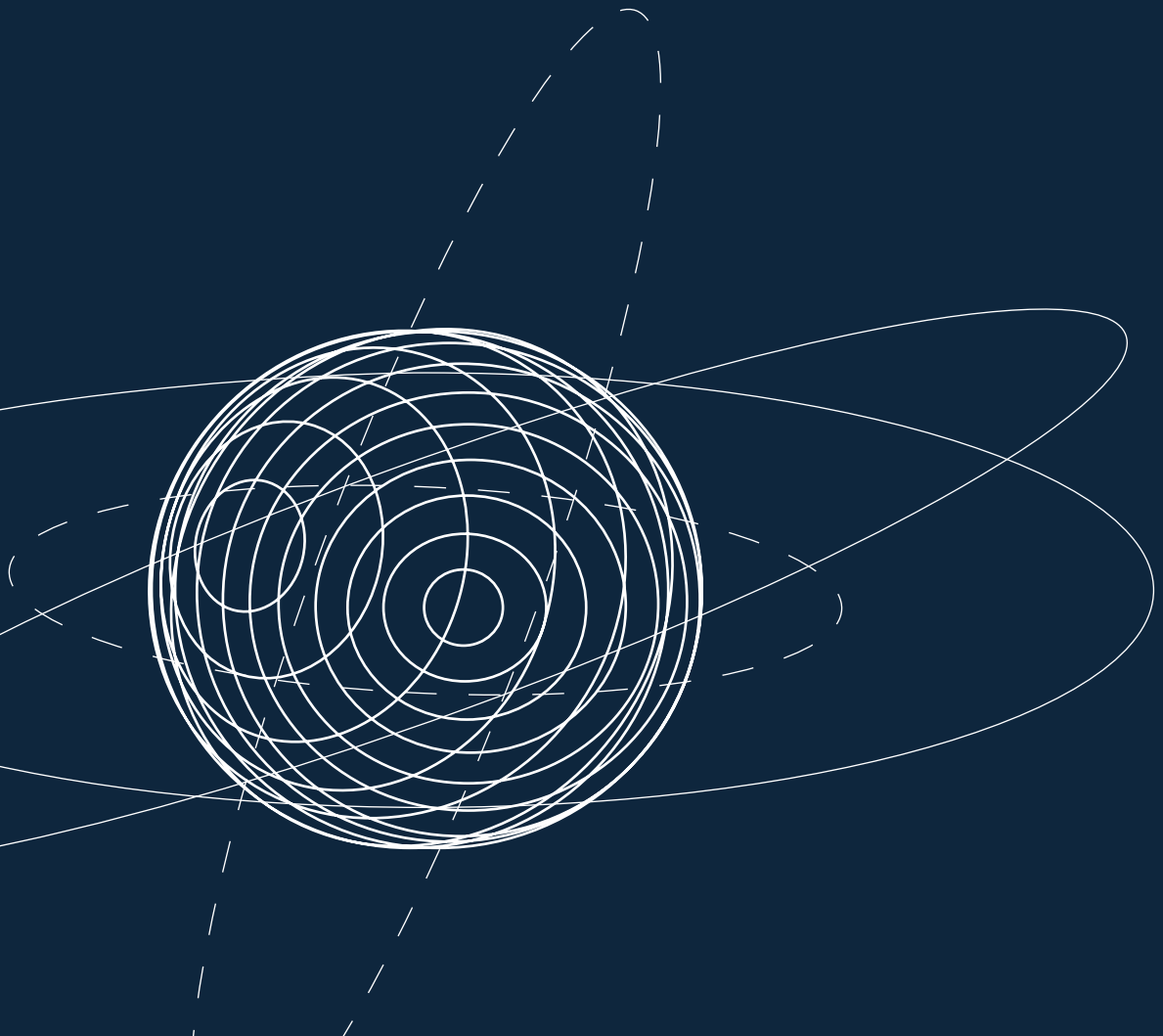
Abbreviations: N.A. = not applicable; DCC = dynamic clinical checklist

Question/score (N = number of participants answering the question)		1 Totally disagree	2 Disagree	3 Neutral	4 Agree	5 Totally agree	N.A.
Usability (U)							
U1	Overall, I am satisfied with how easy it is to use DCC. (N=20)	-	5% (n=1)	-	60% (n=12)	35% (n=7)	-
U2	It was simple to use DCC. (N=20)	-	5% (n=1)	-	55% (n=11)	40% (n=8)	-
U3	I was able to complete the tasks and scenarios quickly using DCC. (N=20)	-	-	20% (n=4)	65% (n=13)	15% (n=3)	-
U4	I was able to efficiently complete the tasks and scenarios using DCC. (N=20)	-	-	15% (n=3)	70% (n=14)	15% (n=3)	-
U5	I felt comfortable using DCC. (N=20)	-	10% (n=2)	5% (n=1)	60% (n=12)	25% (n=5)	-
U6	I quickly understood on how to interact with DCC. (N=20)	-	-	-	45% (n=9)	55% (n=11)	-
U7	It was easy to understand the advices given by DCC. (N=20)	-	-	-	65% (n=13)	35% (n=7)	-
U8	It was easy to find the information I needed. (N=20)	-	-	20% (n=4)	55% (n=11)	25% (n=5)	-
U9	The interface of DCC was pleasant. (N=20)	-	10% (n=2)	15% (n=3)	55% (n=11)	20% (n=4)	-
U10	I liked using the interface of DCC. (N=20)	-	5% (n=1)	25% (n=5)	50% (n=10)	20% (n=4)	-
U11	I could effectively complete the tasks and scenarios using DCC. (N=20)	-	-	-	80% (n=16)	20% (n=4)	-
U12	Whenever I made a mistake using DCC, I could recover easily and quickly. (N=19)	-	-	45% (n=9)	40% (n=8)	10% (n=2)	-
U13	The lay-out of information on the screens was clear. (N=20)	-	20% (n=4)	20% (n=4)	55% (n=11)	5% (n=1)	-

	Question/score	1 Totally disagree	2 Disagree	3 Neutral	4 Agree	5 Totally agree	N.A.
Training & Support (T)							
T1	Training in the use of DCC was sufficient. (N=20)	-	-	15% (n=3)	60% (n=12)	25% (n=5)	-
T2	It was easy to get acquainted using DCC. The manual of DCC was clear. (N=20)	-	-	5% (n=1)	60% (n=12)	35% (n=7)	-
T3	It was easy to find guideline-related information in DCC. (N=20)	-	-	25% (n=5)	50% (n=10)	25% (n=5)	-
T4	The information bullets with guideline-related information were valuable in addition to the checkable items. (N=20)	-	-	10% (n=2)	55% (n=11)	35% (n=7)	-
Behaviour change (B)							
B1	Working with DCC has changed my way of entering patient data. (N=20)	-	10% (n=2)	55% (n=11)	35% (n=7)	-	-
B2	Working with DCC makes me more aware on how to use patient data. (N=20)	-	10% (n=2)	30% (n=6)	55% (n=11)	5% (n=1)	-
B3	Working with DCC has limited the amount of entered patient data. (N=20)	-	25% (n=5)	45% (n=9)	15% (n=3)	15% (n=3)	-
B4	By using the ICU checklist I think I will get less feedback from the pharmacist. (N=20)	5% (n=1)	-	20% (n=4)	45% (n=9)	30% (n=6)	-
B5	I prefer feedback before my actions rather than reminders afterwards. (N=20)	-	-	5% (n=1)	65% (n=13)	30% (n=6)	-
B6	I am prepared to encode patient information in EZIS for use in DCC. (N=20)	-	-	10% (n=2)	70% (n=14)	20% (n=4)	-
B7	By using DCC, I don't spend more time on ICU ward round. (N=20)	-	-	20% (n=4)	55% (n=11)	25% (n=5)	-
B8	The ward round becomes more structured when DCC is used. (N=19)	-	5% (n=1)	5% (n=1)	50% (n=10)	35% (n=7)	-

		1	2	3	4	5	
Question/score		**Totally disagree**	**Disagree**	**Neutral**	**Agree**	**Totally agree**	**N.A.**
Usefulness (Us)							
Us1	I support the use of decision support systems in the ICU. (N=20)	-	-	5% (n=1)	60% (n=12)	35% (n=7)	-
Us2	I like to see DCC-like systems implemented in other departments. (N=20)	-	5% (n=1)	40% (n=8)	35% (n=7)	20% (n=4)	-
Us3	DCC is usable as a training tool. The patient will benefit from DCC. (N=20)	-	-	-	70% (n=14)	30% (n=6)	-
Us4	I think that the ICU ward round checklist of DCC can prevent medical errors. (N=20)	-	-	5% (n=1)	50% (n=10)	45% (n=9)	-
Us5	I think DCC can improve the quality of care on the hospital wards. (N=20)	-	-	-	80% (n=16)	20% (n=4)	-
Us6	If the ICU ward round checklist of DCC is not available I have the feeling of forgetting items. (N=20)	-	10% (n=2)	45% (n=9)	35% (n=7)	10% (n=2)	-
User satisfaction (G)							
G1	DCC generates the right amount of checkable items for the ICU ward round checklist. (N=20)	-	10% (n=2)	15% (n=3)	65% (n=13)	10% (n=2)	-
G2	Overall, I think DCC is a useful tool. (N=20)	-	-	-	80% (n=16)	20% (n=4)	-
G3	It is convenient that DCC can automatically check items based on medical rules. (N=20)	-	15% (n=3)	5% (n=1)	60% (n=12)	20% (n=4)	-
G4	DCC generates correct checkable items for most patients. (N=19)			5% (n=1)	65% (n=13)	30% (n=6)	-
G5	Overall, I am satisfied with DCC. (N=20)	-	-	5% (n=1)	65% (n=13)	30% (n=6)	-
G6	I think the concept of a DCC has potential. (N=20)	-	-	-	70% (n=14)	30% (n=6)	-
G7	I rate the ICU ward round checklist of DCC with a (1 to 5): (N=20)	**4.13 out of 5** (95% confidence interval of 3.91 to 4.34)					

4.2



Checklists, cognitive aids, and the future of patient safety

C. S. Webster

Accompanying editorial in British journal of anaesthesia 2017

On Wednesday, October 30, 1935, an evaluation flight of the Boeing Model 299 was undertaken at Wright Field, northeast of Dayton, OH, USA. The Model 299 was the most technologically sophisticated aircraft of its time and was nicknamed the Flying Fortress because of the extent of its armaments. Major Ployer P. Hill was the pilot, and it was his first flight in the new aircraft. The aircraft appeared to ascend normally, but suddenly stalled, turned on one wing, and crashed, killing two of the aircraft's five crew, including Major Hill. The investigation into the crash discovered that Major Hill had omitted a crucial step during the preflight preparation; he forgot to release a catch, which on the ground locked the aircraft's control flaps.[1] Once in the air, this mistake rendered the aircraft uncontrollable. The crash investigators knew that there was probably no one better qualified to fly the new aircraft than Major Hill—his co-pilot was also highly qualified—yet despite this, the fatal error was still made. The investigators concluded that given the experience of the pilots, further training would not be an effective response to prevent such an event from happening again; a response that is very different from that which often occurs in health care when a mistake is made.[2] Some commentators initially believed that this meant the new aircraft was simply too complicated to fly reliably. A new approach was needed, and it took the form of a simple list of crucial tasks that must be completed before the aircraft could leave the ground. The first aviation checklist had been devised.[1] With the checklist in use, despite the aircraft's sophistication, the Model 299 (and later versions of it) performed safely for many years.

Around 70 yr later, the crash of the Model 299 and creation of the aviation checklist were the inspiration for the development of the now celebrated World Health Organization (WHO) Surgical Safety Checklist.[1] The technical issues for surgical safety were similar to those in aviation; highly qualified and skilled clinicians working in the high-technology environment of the operating room needed to ensure that certain crucial steps were not omitted during a procedure. The WHO Surgical Safety Checklist was therefore designed to improve team communication and consistency of care by prompting checking and communication at crucial points. In a large-scale multinational study of 7688 patients reported in 2009, use of the WHO Surgical Safety Checklist was shown to reduce the overall rate of postoperative complications by 36%.[3] In the succeeding years, there have been a flurry of safety checklist studies, which have included the emergence of a better understanding of the limitations of the use of checklists in surgery and health care.[4–7]

One substantial limitation of applying aviation-type checklists in health care is the fact that although aircraft are complicated, patients undergoing health care are complex.[2,8,9] The challenge of patient variability should not be underestimated. Unlike many high-technology endeavours where a great deal of standardization is possible, health care clearly must contend with the subtle physical variations and abnormal anatomies

and pathologies that exist in individuals; differences that are often unknown and unknowable before the procedure has begun. This represents a different situation from that with a machine, such as an aircraft, where its exact structure and function is known and where these details are documented. Checklist design for aircraft, where the vast majority of eventualities can be anticipated, is therefore a relatively simpler task than attempting to adopt the same approach in health care.

However, despite such limitations, systematic reviews of the use of safety checklists in the operating room demonstrate their substantial benefits in terms of improving patient outcomes, but only when teams engage with the checklist process and when compliance with checklist items is high.[10–15] One study found no improvements in postoperative survival rates when checklists were not completed or when completed only in part, but showed significant survival benefits when checklists were fully completed.[16] Checklist design is not a trivial process. The checklist should be short; its design must be based on the best clinical knowledge, and it must not be influenced by managerial concerns regarding the medico-legal protection of the organization.[1, 17, 19] A formal process for the introduction of a safety checklist is typically needed so that clinicians know how the checklist should be used.[4, 7] Engagement by key team personnel is also important to establish a safety culture that encourages and maintains compliance with the checklist for every patient.[5, 18]

The article by De Bie and colleagues[20] in this issue of the British Journal of Anaesthesia describes an in situ simulation study of a new electronic dynamic clinical checklist (DCC), which contains two significant innovations with the potential to solve a number of important problems in the successful use of checklists in health care and to advance patient safety more widely.[21, 22] These innovations are as follows: (1) meaningful sharing and integration of information between multiple hospital systems; and (2) automatic preparation of a personalized electronic checklist of items relevant to the care of each individual patient. The DCC system achieves this by using a set of algorithms to select checklist items relevant to each patient in the intensive care unit based on information accessed from the patient's electronic health record, the hospital's treatment protocols, and pharmaceutical databases. The algorithms can also automatically check certain items when the system has access to the relevant information, hence reducing the checklist burden on the clinician. Comparing the use of their hospital's standard paper-based checklist with the new DCC during 116 in situ simulations demonstrated an increase in completion rate of checklist items from 74 to 100%. Participants rated their satisfaction with the DCC highly and agreed that the approach had potential in medical care. In addition, follow-up by the pharmacist after the simulated ward round, as prompted by alerts from the hospital's clinical decision support system, reduced dramatically from

80% to only 3.6% with use of the DCC. The use of simulation is becoming more common for the purposes of evaluating new safety interventions and in making inferences about team behaviour in the clinical setting.[23] Given the evidence that compliance with checklists is an essential part of their effectiveness in improving patient outcomes, we might therefore expect the DCC to have substantial potential to improve clinical care in the intensive care unit, and I look forward to these clinical studies.

Many hospital systems and devices currently have some facility for sharing certain information with other devices, but few have achieved the kind of meaningful, safety-orientated integration that is reported here with the DCC. One potential risk of the success of the WHO Surgical Safety Checklist is that the use of checklists has now become so widely mandated throughout health care that poorer quality checklists may be introduced into use, and checklists may be introduced into practice areas where they are less effective; both outcomes are likely to lead to disengagement by clinicians.[9, 24, 25, 26] In contrast, allowing the algorithms of the DCC to access all relevant data when generating checklist items for individual patients means that the resulting personalized checklist is immediately relevant to the patient's care. Unlike a paper-based or static checklist, non-relevant or generic checklist items need not appear on the DCC. From a psychological perspective, the salience of any message or signal is determined by its informational content or informativeness, hence messages that contain misinformation or false alarms tend quickly to be ignored.[27] Therefore, a checklist with few or no generic items would be expected to be more salient for the user. As the authors state, in this sense the DCC is a true cognitive aid, in that it supports and assists the clinician in getting his or her job done, rather than potentially being viewed as a mandatory requirement, of variable relevance, that might add further burden to their existing workload. The DCC is therefore likely to engage clinicians better and to encourage them to check every item, as occurs during every flight with an aviation checklist. Further research considering what happens to clinicians' work patterns when the DCC is used in the clinical setting and whether it has indeed become integrated into their workflow will be interesting, particularly given that conversion of other formerly physical records into electronic formats (e.g. radiographs and patient notes) has often had unanticipated consequences.[22]

I was interested that the feature which allows certain checklist items to be completed automatically by the DCC could be overridden by clinicians, if they preferred to complete such a check themselves. The tailoring of the set of algorithms of such a dynamic checklist system is clearly important for many reasons; in order to adjust sensitivity to the kinds of events that clinicians want to monitor, to update the checklist items when clinical knowledge changes, and to customize the checks for particular patient populations or

clinician preferences. If systems such as the DCC become more widespread, I expect that additional work will be done to fine-tune the algorithms that generate the checklists. This work could determine what kinds of information the checklist algorithms need to access to make the best checklists, and what the optimal hierarchy or prioritization of checklist items might be to produce a checklist that tells you all you need to know but isn't too long. Electronic systems, such as the DCC, make it easy to update such features, because like all software, updates can propagate out from a central location to all devices in the network, and there will be no physical copies of the old version of the checklist to remove from use.

The DCC represents an example of a system where electronic clinical information has been meaningfully synthesized from various hospital systems, and non-relevant information has been filtered out. I believe such an approach will have many applications in the improvement of the quality and safety of patient care in the near future, particularly if we are indeed at the dawn of medicine's computer age.[21, 22, 25]

One pressing area of need for such an approach is that of alarm management in operating rooms and intensive care units, and this is an area where health care could again benefit from the techniques used in aviation. The functional integration possible in many clinical devices is currently limited and hampered by various different proprietary formats and standards. The practical consequence of this is that many devices, from drug infusion pumps to patient monitors, generate their own stream of alerts and alarms independently of each other, without any co-ordination or prioritization, leading to a cacophony of auditory alerts where important alarms can be lost amongst trivial ones. This leads to alarm fatigue, where alarms may be ignored or switched off. A recent study of this problem reported from a single hospital, with 77 intensive care beds, recorded the occurrence of an astonishing 2.558.760 unique physiological alarms during intensive care in a single month.[28] In aviation, the alarm fatigue problem is managed by engineers and pilots working co-operatively to agree upon exactly what needs to be alerted to the pilot from all aircraft systems and what does not. Agreed alarms are then placed in a hierarchy, with many events being reported only as 'cautions' or 'advisories' on a screen, but without any auditory alert. Pilots would not tolerate the alarm chaos that clinicians currently face. Even an event as apparently serious as an engine failure in a multi-engine aircraft will not result in a top-level alarm with an auditory alert, but only a caution. This is because such an event does not require immediate pilot intervention owing to the automatic systems on modern aircraft.[22] The manufacturers of components for aircraft cockpits must meet very specific compatibility standards, but at present this is not the case in health care. Although checklists, either dynamic or otherwise, are a successful

approach to align and increase the consistency of key procedural aspects of patient care, such alignment needs to extend beyond procedures to include the equipment used in clinical environments.[29]

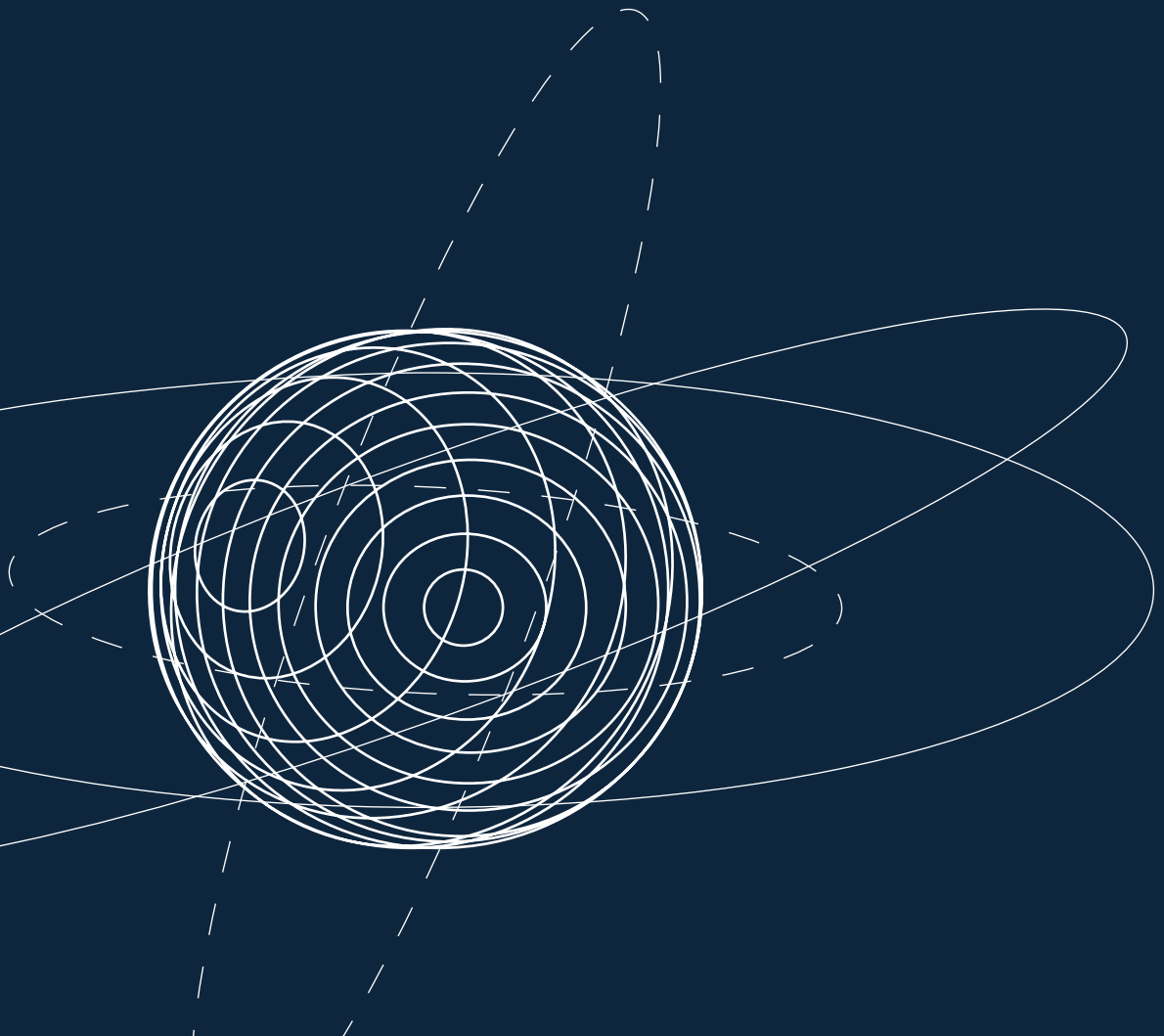
We know that paper checklists, when well designed, properly introduced, and complied with, can substantially reduce the burden of postoperative complications. The electronic DCC reported in this issue of the British Journal of Anaesthesia represents an important development beyond paper or static checklists, in that the checklist is automatically tailored to each patient by drawing on various sources of patient data. The results of a simulation study in the intensive care unit are encouraging, including excellent checklist compliance. The next step will be clinical trials of the DCC in order to determine whether the excellent compliance rates seen in the simulator translate into improvements in the safety and quality of patient care. The information filtering and prioritization features of a dynamic checklist also seem highly suitable for solving other difficult problems in health care, such as the alarm management problem.

REFERENCES

1. Gawande A. *The Checklist Manifesto — How to Get Things Right*. New York: Metropolitan Books, 2009
2. Webster CS, Anderson BJ, Stabile MJ, et al. Improving the safety of pediatric sedation: human error, technology and clinical microsystems. *Pediatric Sedation Outside of the Operating Room: a Multispecialty International Collaboration*. New York: Springer Science, 2015; 587–612
3. Haynes AB, Weiser TG, Berry WR, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2009; **360**: 491–9
4. Urbach DR, Govindarajan A, Saskin R, et al. Introduction of surgical safety checklists in Ontario, Canada. N Engl J Med 2014; **370**: 1029–38
5. Singer SJ, Molina G, Li Z, et al. Relationship between operating room teamwork, contextual factors, and safety checklist performance. J Am Coll Surg 2016; **223**: 568–80
6. de Vries EN, Prins HA, Crolla RMPH, et al. Effect of a comprehensive surgical safety system on patient outcomes. N Engl J Med 2010; **363**: 1928–37
7. Marshall SD. Helping experts and expert teams perform under duress: an agenda for cognitive aid research. Anaesthesia 2017; **72**: 283–95
8. Webster CS. Why anaesthetising a patient is more prone to failure than flying a plane. Anaesthesia 2002; **57**: 819–20
9. St Pierre M, Hofinger G, Simon R. Crisis Management in Acute Care Settings—Human Factors and Team Psychology in a High-Stakes Environment, 3rd Edn. Cham, Switzerland: Springer International 2016
10. Treadwell JR, Lucas S, Tsou AY. Surgical checklists: a systematic review of impacts and implementation. BMJ Qual Saf 2013; **23**: 299–318
11. Thomassen Ø, Storesund A, Søfteland E, et al. The effects of safety checklists in medicine: a systematic review. Acta Anaesthesiol Scand 2014; **58**: 5–18
12. Borchard A, Schwappach DLB, Barbir A, et al. A systematic review of the effectiveness, compliance, and critical factors for implementation of safety checklists in surgery. Ann Surg 2012; **256**: 925–33
13. Russ S, Rout S, Sevdalis N, et al. Do safety checklists improve teamwork and communication in the operating room? A systematic review. Ann Surg 2013; **258**: 856–71
14. Bergs J, Hellings J, Cleemput I, et al. Systematic review and meta-analysis of the effects of the World Health Organization surgical safety checklist on postoperative complications. Br J Surg 2014; **101**: 150–8
15. Jammer I, Ahmad T, Aldecoa C, et al. Point prevalence of surgical checklist use in Europe: relationship with hospital mortality. Br J Anaesth 2015; **114**: 801–7
16. van Klei WA, Hoff RG, van Aarnhem EEHL, et al. Effects of the introduction of the WHO Surgical Safety Checklist on inhospital mortality: a cohort study. Ann Surg 2012; **255**: 44–9
17. Webster CS, Mason KP, Shafer SL. Threats to safety during sedation outside of the operating room and the death of Michael Jackson. Curr Opin Anesthesiol 2016; **29**: S36–47

18. Weiser TG, Berry WR. Review article: perioperative checklist methodologies. Can J Anesth 2013; **60**: 136–42
19. Hill MR, Roberts MJ, Alderson ML, Gale TC. Safety culture and the 5 steps to safer surgery: an intervention study. Br J Anaesth 2015; **114**: 958–62
20. De Bie AJR, Nan S, Vermeulen LRE, et al. Intelligent dynamic clinical checklists improved checklist compliance in the intensive care unit. Br J Anaesth 2017; (in this issue)
21. Webster CS. Health care technology, the human–machine interface, and patient safety during intravenous anesthesia. Total Intravenous Anesthesia and Target Controlled Infusions: a Comprehensive Global Anthology. Cham, Switzerland: Springer International 2017; 667–83
22. Wachter R. The Digital Doctor—Hope, Hype, and Harm at the Dawn of Medicine's \ Computer Age. New York: McGraw Hill, 2015
23. Merry AF, Hannam JA, Webster CS, et al. Retesting the hypothesis of a clinical randomized controlled trial in a simulation environment to validate anesthesia simulation in error research (the VASER study). Anesthesiology 2017; **126**: 472–81
24. Webster CS, Andersson E, Edwards K, et al. Deviation from accepted drug administration guidelines during anaesthesia in twenty highly realistic simulated cases. Anaesth Intensive Care 2015; **43**: 698–706
25. Simpao AF, Ahumada LM, Rehman MA. Big data and visual analytics in anaesthesia and health care. Br J Anaesth 2015; **115**: 350–6
26. Ko HCH, Turner TJ, Finnigan MA. Systematic review of safety checklists for use by medical care teams in acute hospital settings—limited evidence of effectiveness. BMC Health Services Research 2011; **11**: 211. Available from http://www.biomedcentral.com/1472-6963/11/211 (accessed 15 June 2017)
27. Li B, Parmentier FBR, Zhang M. Behavioral distraction by auditory deviance is mediated by the sound's informational value. Exp Psychol 2013; **60**: 260–8
28. Drew BJ, Harris P, Ze`gre-Hemsey JK, et al. Insights into the problem of alarm fatigue with physiologic monitor devices: a comprehensive observational study of consecutive intensive care unit patients. PLoS One 2014; **9**: e110274
29. Pronovost PJ, Ravitz AD, Stoll RA, et al. Transforming Patient Safety—a sector-wide systems approach. Report of the WISH Patient Safety Forum, 2015. Available from http://www.wish-qatar.org/app/media/1430 (accessed 15 June 2017)

4.3



Intelligent checklists improve checklist compliance in the intensive care unit: a prospective before-and-after mixed-method study

A.J.R. De Bie, E.H.J. Mestrom, W. Compagner, S. Nan, L. van Genugten, K. Dellimore, J. Eerden, S. van Leeuwen, H. van de Pol, F. Schuling, X. Lu, A.J.G.H. Bindels, R.A. Bouwman, H.H.M. Korsten

British journal of anaesthesia 2020

ABSTRACT

Background

To determine whether an intelligent dynamic checklist increases the compliance to best eligible practice compared to a paper checklist during ward rounds on an Intensive Care Unit (ICU).

Methods

A single-centre prospective before-and-after mixed-method trial in a 35-bed mixed medical and surgical ICU. Daily ICU ward rounds were observed during two periods of eight weeks: a control period with paper checklists, and an intervention period in which the intelligent checklist was available. The primary outcome was compliance with best eligible practice, measured as the percentages of checked items and unchecked critical items. Secondary outcomes included patient-related outcomes, violation of guidelines, and the usability of intelligent checklists.

Results

36 clinicians were observed while visiting 197 patients during 352 rounds in the control period versus 211 patients during 366 rounds in the intervention period. The patients' baseline characteristics and severity of disease were well balanced. In the intervention period more items were checked compared to the control period (interquartile range 94.4–100.0) vs 75.1% (66.7–86.4); z=21.9, p=0.03), less critical items remained unchecked (median 0.0 (0.0–0.0) vs 15.4% (8.3–27.3); z=-17.7, p=0.01), and the ICU length of stay was reduced (median 1 (1-3) vs 2 (1-4) days; z=-2.55, p=0.05). Although the usability needs to be improved, clinicians perceived the intelligent checklist as an innovative suitable tool to replace paper checklists.

Conclusion

The availability of an intelligent checklist during ICU ward rounds improves compliance to best eligible practice compared to a paper checklist on a mixed ICU.

INTRODUCTION

Implementation of medical checklists in clinical practice reduces adverse events and related deaths.[1-5] However, not all subsequent studies could reproduce these beneficial effects.[5-9] This might be explained by a lack of compliance due to multiple socio-organizational barriers, as well as checklist factors, including the checklist design, accessibility, workflow integration and perceived relevance of content.[2, 6, 8] Digital checklists have the potential to meet the necessary requirements to overcome these barriers, to translate medical knowledge and evidence to the bedside, and optimize the compliance.[10-13]

Recently, a novel clinical decision support system (CDSS) called TraceBook has been developed that uses dynamic clinical checklists (DCC) in a process-oriented and context-aware manner.[14, 15] This CDSS contains several innovations to support successful use of checklists in healthcare. First, the CDSS is able to gather and integrate information from different data sources within the hospital. Second, the rule engine within the CDSS prepares personalized digital checklists containing items relevant to the care of each individual patient.[11, 12, 15] Third, automated checks are feasible when healthcare professionals locally agree that a rule can be checked automatically. Finally, the CDSS provides better insight into workflow, displays guideline recommendations upon request, and highlights relevant data from the medical databases requiring extra attention such as laboratory results.[14, 15]

The results of a simulation study in 2017 showed that implementing the DCC for Intensive Care Unit (ICU) ward rounds was associated with improved compliance to local guidelines compared to local standard of care with a paper checklist (73 vs 100%).[11] Participating physicians appreciated the DCC with a high satisfaction score (4.1 out of 5).[11] Although promising, these results shed no light on compliance and effectiveness of the DCC in real clinical practice. Therefore, we conducted this before-and-after mixed-method trial to evaluate and provide context about the DCC's effect on the compliance with best eligible practice during ICU ward rounds compared to the local standard of care using paper checklists.

METHODS

The guidelines for Strengthening the Reporting of OBservational studies in Epidemiology (STROBE) were applied to prepare this article.[16]

Study design and setting

This prospective before-and-after mixed-method study was carried out from July 2018 until March 2019 in the ICU of Catharina Hospital Eindhoven, a tertiary hospital in the Netherlands. This department is a 35-bed mixed medical and surgical ICU, including cardiothoracic surgery (department characteristics are described in eTable 1, Supplement 1). The study protocol is available in Supplement 2. The study consisted of two periods: a control period of eight weeks with local standard of care and an intervention period of eight weeks. A mixed-methods design, including questionnaires and interviews, was used to provide context for the quantitative results. The study was approved by the Institutional Review Board (W18.046) and registered at clinicaltrials.gov (NCT03599856).

Study population and eligible criteria

Eligible participants of this study were intensivists, residents, and ICU physician assistants who were in the lead of the ICU daily ward round.17 The intensivist carried the final responsibility during these rounds. All clinical ICU staff were informed about the study and all consented to participate.

Local standard of care

An ICU ward round is a scheduled visit of the ICU patients at the end of the morning in which residents or physician assistants review relevant clinical data and clinical decisions are made together with the responsible intensivist. A bedside paper checklist containing 17 items is available to be used at their convenience. This checklist is based on the FAST HUG mnemonic with extra items added based on local guidelines developed since its introduction in the ICU (Supplement 2).[18, 19] In addition, the hospital uses the CDSS Gaston® (Gaston Medical, Eindhoven, the Netherlands) to review prescribed medication on the ICU and alert pharmacists if one of the 23 predetermined pharmacological clinical rules for the ICU are violated.[20-22]

The intervention: Dynamic Clinical Checklist (DCC) for the ICU ward round

TraceBook's DCC generates on request dedicated checklists for each individual patient and is easily accessible on a tablet or computer (*Figure 1*). To generate the DCC, the systems of TraceBook and GASTON both have a rule engine containing a model of transparent

algorithms, comparable with a decision tree, with clinical rules and pharmacological rules. First GASTON gathers the relevant medical data of a patient from various medical information systems, such as the electronic health record, the laboratory information system, the pharmaceutical prescription system, and others. Then both systems run the rule engines and TraceBook determines which rules are relevant for an individual patient in a specific context. These items then become a checkable item for the DCC of that particular patient (eFigure 1 in Supplement 1 provides a schematic overview of how a DCC is composed).[20, 21, 23] After considering an item users can tab the corresponding box and make a note if desired until the DCC is completed. The whole system is designed to easily create or modify rules, even by caregivers themselves. The model for the ICU ward round DCC is comparable with the DCC used in the previous simulation-based trial.[11] Prior to starting the study, algorithms were updated by researcher ADB and checked by researcher AB to match currently applied local guidelines.[11] Rules were not modified during both periods.

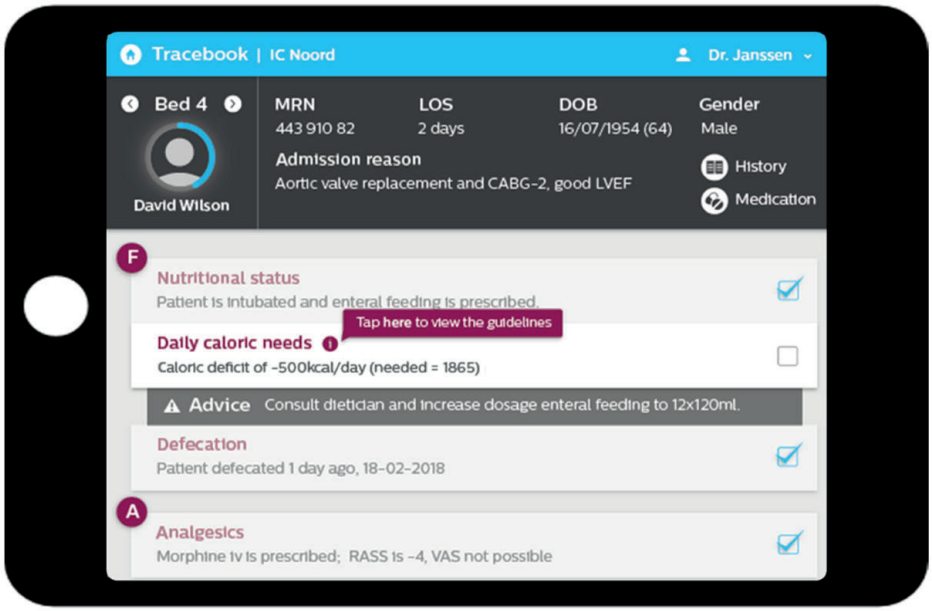


FIGURE 1. Screenshot of TraceBook's dynamic clinical checklists.
Screenshot of TraceBook's dynamic clinical checklists for the ICU ward round from a fictional person.

Data collection and endpoints

Two researchers (ADB and EM) observed the rounds. Both observers were former ICU residents and familiar with the local practice and guidelines. A ceiling-mounted camera and a microphone allowed the researchers to observe the rounds in another room out of the sight of the ICU staff (video was not recorded). Except for the weekends, all morning rounds were eligible to be observed if observers were available (eFigure 2, Supplement 1).

The primary outcome is compliance with best eligible practice assessed as the percentage of discussed checklist items and the percentage of critical items that remained unnoticed on each ward round of a patient admitted to the ICU. A critical item was defined as an item that required an intervention based on local protocol. A standardized paper list with 25 predefined items, based on the paper checklist and local guidelines (eTable 2, Supplement 1), combined with the output of the DCC were used to judge which items needed to be discussed (e.g. the item "radiological examination" was considered inapplicable if no examination was performed in the last 24 hours).

Secondary outcomes were mortality rates, length of stay and number of ventilator days. Other secondary outcomes were the number of automatically checked items and the following care processes: the number of violated pharmacological clinical rules and registered complications, the number of days with prescribed regular use of analgesics, sedatives and empiric antibiotics, the pain scores (Critical Care Pain Observation Tool (CPOT;0 – 6) and Visual Analogue Scale (VAS;0 – 10)),[23] and the Richmond Agitation-Sedation Scale (RASS;-5 – 5).[24]

User experience was evaluated using self-report questionnaires and semi-structured interviews at the end of each period. The AttrakDiff questionnaire was used to assess usability based on pragmatic and ease-of-use (hedonic) factors (Supplement 2).[25] User acceptance was assessed with a questionnaire based on the Technology Acceptance Model-2 (TAM-2; Supplement 2).[26] To better understand the quantitative insights on user acceptance with DCC, interviews were conducted. An independent researcher (LG) conducted semi-structured interviews in the two weeks after the intervention period. An interview topic guide based on the TAM-2 model was used to explore acceptance, perceived strengths and weaknesses of the DCC, expectations and experiences, and the perceived barriers to implementation (eTable 7, Supplement 1).

Statistical analyses

A sample size calculation was performed with g*power (g*power team, version 3.1.9.2, Kiel, Germany). Based on findings of the pilot study (73.6% (IQR:64.5 – 79.3) vs 100%

(IQR:100.0 – 100.0),[11] a sample of 50 observed patients during each period would provide 95% power to detect a difference of 26.4% of checked items with a type I error of 5% and corrected for dropouts. We aimed for 120 patients in each period, since in contrast to a simulation-based study the patient scenarios in real practice are not controlled by the researchers.

Quantitative data analyses were performed with SPSS (version 22.0; IBM Corp, Armonk, NY, USA). Distribution of continuous variables was assessed with the Kolmogorov-Smirnov test and by analyses of the histograms. The not normally distributed data were analysed with the Mann-Whitney U test or the Chi-square test. A false discovery rate correction was used to correct for the multiple comparisons and calculate the false discovery rate adjusted p-values.27 All the reported p-values are two-sided, all have been adjusted, and a p-value of 0.05 or lower was considered statistically significant..

A deductive approach was applied for the qualitative data analysis, with categories based on the TAM-2 with additional elements around the topic of routines and habits. Analyses started with annotations at the sentence, question and topic level on the interview transcripts by two independent researchers (LG and KD) with Atlas.ti 7 (Atlas.ti, Scientific Software Development GmbH, Berlin, Germany, 2013). For each factor, variations in opinions within the factors were described (e.g. positive vs negative; pros vs cons).

RESULTS

The participants and daily ward rounds

In both periods 14 intensivists, 7 ICU physician assistants, and 15 residents were observed during their daily ICU rounds. In the intervention period three new residents substituted eight residents that worked on the ICU during the control period due to their rotating internships. In the control period from July through August 2018, 196 individual patients were observed in 352 rounds. From September through November 205 patients were observed in 366 rounds in the intervention period. Baseline characteristics, severity-of-disease classification score and comorbidities were well balanced between both periods (eTable 3, Supplement 1). Both the histograms and the tests for normality indicated that the data were not normally distributed (Appendix 1, Supplement 1).

Primary outcome

Figure 2 illustrates an increase of the median percentage of checked overall items from 77.8% (IQR = 66.7 – 86.4) in the control period to 100% (IQR = 94.4 – 100.0) in the

intervention period ((p=0.03, z=-22.3; *Table 1*). The median percentage of unchecked critical items decreased (p=0.02, z=-16.2; *Table 1*) from 15.4% (IQR = 8.3 – 27.3) to 0.0% (IQR = 0.0 – 0.0). The false discovery rate correction did not change these findings (eTable 4, Supplement 1).

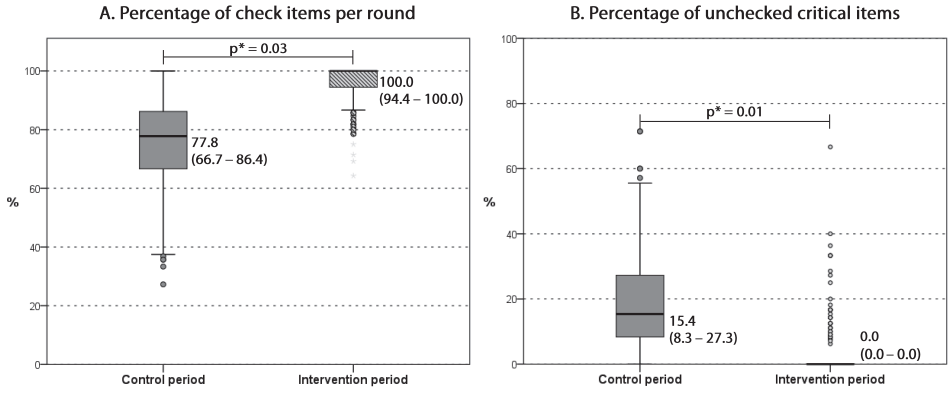


Figure 2. Boxplots of the checked items and unchecked critical items for each group.
Boxplots of the median percentage checked items (A) and unchecked critical items (B) per ICU ward round (= adjusted p-value).*

Secondary outcomes

Patient-centred outcomes

The length of stay in the ICU (2.0 (1.0 – 4.0) vs 1.0 (1.0 – 4.0) days; z-score -2.5, p=0.05) and hospital (9.0 (6.8 – 17.0) vs 8.0 (6.0 – 16.0) days, z-score -2.5, p=0.05) was shorter during the intervention period (*Table 2*). Mortality rates and invasive ventilation time were similar in both groups (*Table 2*). The difference of the ICU length of stay remained significant after the false discovery rate correction (eTable 4, Supplement 1).

Outcomes related to specific care processes

The median CPOT score was lower during the intervention period compared to the control period, while the median number of days with intravenous sedatives prescribed was higher. The number of days in which opiates were prescribed for regular use and empiric antibiotics was reduced in the intervention period (*Table 1*). Other secondary outcomes did not differ between both periods (*Table 1*). The false discovery rate correction did not change these finding, except for the difference in use of opiates which became not significant (p=0.17, eTable 4, Supplement 1). Post hoc correlation

Table 1 – Primary and secondary outcomes

	Control period Paper checklist (*n* = 352)	Intervention period Digital dynamic checklist (*n* = 366)	χ^2- or z-score	*p value*[#]
Primary outcomes				
Percentage of checked items, %: Median (IQR)	77.8 (66.7 – 86.4)	100.0 (94.4 – 100.0)	z = 21.9	0.03
Percentage of unchecked critical items, %: Median (IQR)	15.4 (8.3 – 27.3)	0.0 (0.0 – 0.0)	z= -17.7	0.02
Secondary outcomes				
Medication related rules				
Number of alerts, Median (min – max)	0 (0 – 32)	0 (0 – 14)	z = -1.52	0.30
Number of relevant alerts*, Median (min – max)	0 (0 – 6)	0 (0 – 3)	z = -0.15	0.98
Phone calls of pharmacist to ICU clinician, Median (min – max)	0 (0 – 5)	0 (0 – 3)	z = -0.15	1.02
Number of interventions based alerts, Median (min – max)	0 (0 – 3)	0 (0 – 3)	z = -0.21	1.08
Prescribed medication: Number of days per patient; Median (min – max)				
Opiates; prescribed as regular use each day	1.0 (0.0 – 24.0)	1.0 (0.0 – 14.0)	z = -2.00	0.17
Paracetamol; prescribed as regular use each day	2.0 (0.0 – 47.0)	1.0 (0.0 – 30.0)	z = -1.32	0.41
No PPI, while indicated	1.0 (0.0 – 37.0)	1.0 (0.0 – 30.0)	z = -0.30	1.16
Intravenous sedatives**	0.0 (0.0 – 11.0)	1.0 (0.0 – 12.0)	z = -7.09	0.01
Antibiotics***	0.0 (0.0 – 34.0)	0.0 (0.0 -18.0)	z = -3.78	<0.01
Complications				
Registered complications, median (min – max)	0.0 (0.0 – 19.0)	(0.0 – 11.0)	z = -0.34	1.17
Gastro-intestinal bleedings, n (%)	1 (0.5)	5 (2.4)	χ^2 = 1.88	0.47
Hospital acquired pneumonia, n (%)	8 (4.1)	5 (2.4)	χ^2 = 0.48	0.82
CRBSI, n (%)	3 (1.5)	0 (0.0)	χ^2 = 3.24	0.21
Hypoglycaemia (<4 mmol L^{-1}), median (min – max)	0.0 (0.0 – 9.0)	(0.0 – 6.0)	z = -0.19	1.02
Hyperglycaemia (>15 mmol L^{-1}), median (min – max)	0.0 (0.0 – 9.0)	(0.0 – 5.0)	z = -1.85	0.19
Days without defecation for at least >48hours, median (min – max)	1.0 (0.0 – 7.0)	0.0 (0.0 – 5.0)	z = -1.61	0.28

Table 1 – Continued

	Control period Paper checklist (*n* = 352)	**Intervention period Digital dynamic checklist (*n* = 366)**	**χ- or z-score**	***p value*#**
VAS (1-10), n	n=1052	n=1266		
Median (IQR)	1.0 (0.0 – 3.0)	1.0 (0.0 – 3.0)	z = -0.91	0.65
Pain scores (1-10) >4; n (%)	177 (16.8)	207 (16.4)	χ^2 = 0.06	1.09
CPOT (0-6), n	n=453	n=404		
Median (IQR)	1.0 (0.0 – 2.0)	0.0 (0.0 – 1.8)	z = -4.98	<0.01
CPOT score (0-6) >2; n (%)	115 (25.4)	64 (15.8)	χ^2 = 11.20	<0.01
RASS (-5 – 5), n	n=1158	n=1270		
Median (IQR)	0.0 (-1.0 – 0.0)	0.0 (-1.0 – 0.0)	z = -1.22	0.44

Abbreviations: IQR = interquartile range; CRBSI = central-venous-catheter-related bloodstream infections; PPI = proton pump inhibitor

determined by the hospital pharmacist on duty; ** Propofol or Midazolam; * Days with only selective digestive decontamination excluded;*

plots suggest that in the control period more critical items remained unnoticed in patients with a higher APACHE IV score at admission, while in the intervention period a balanced or even smaller negative relationship was found (eFigure 3, Supplement 1).

Table 2 – Clinical outcomes of the patients in the control and the intervention period

	Control period Paper checklist (*n* = 197)	**Intervention period Digital dynamic checklist (*n* = 211)**	**χ^2- or z-score**	***p value*[#]**
Mortality, n (%)				
ICU,	12 (6.1)	14 (6.6)	χ^2 = <0.01	1.01
30-day	17 (8.6)	20 (9.5)	χ^2 = 0.02	0.96
90-day	23 (11.7)	28 (13.3)	χ^2 = 0.11	1.06
Length of stay in days, median IQR)				
ICU	2.0 (1.0 – 4.0)	1.0 (1.0 – 3.0)	z = -2.55	0.05
Hospital	9.0 (6.8 – 17.0)	8.0 (6.0 – 16.0)	z = -2.46	0.05
Invasive ventilation time in hours, median (IQR)				
Overall group	7.0 (4.0 – 23.3)	7.0 (4.0 – 30.0)	z = -0.02	0.98
Patients >24hrs of invasive ventilation	89.0 (42.0 – 117.0)	68.0 (42.8 – 173.0)	z = -0.20	1.05

[#]*False Discovery Rate adjusted p-value*

Descriptive summaries

Control period

On an average a participant performed 4 ICU rounds (range 1 – 8) that were eligible for observation per day. In each round was a paper checklist available with on average 14 items being applicable to the patient (range 8 – 22). In total, 3764 items (75.2%) of the 5007 items that were applicable to the patients were discussed by the participants during the 352 rounds. Of all the critical items remained 18.8% (629 of 3351) unnoticed during these 352 rounds. These 629 critical items included 306 violated pharmacological clinical rules. In 4 of the 352 rounds (1.1%), were all applicable checklist items discussed by the end of the ward round, while in 75 of the 352 rounds (21.3%) no critical items remained unnoticed. The average percentage of completed checklist items and unnoticed critical items per day were 74.8% (range 58.8 – 89.8) and 18.6% (range 7.9 – 34.8), respectively.

Intervention period

On an average a participant performed 5 ICU rounds (range 3 – 8) that were eligible for observation per day. In each round the DCC was available which contained on average 10 applicable items per checklist (range 5 – 18). In total, 5332 items (97.4%) of the 5476 applicable items was checked by the participants during the 366 rounds. Of all the critical items remained 2.0% (64 of 3198) unchecked during these 366 rounds. The DCC checked 1890 of the 5476 applicable checklist items automatically, which reduced the number of checklist items with an average of 5 items per checklist (range 0 – 11). In 263 of the 366 rounds (61.3%) were all applicable checklist items checked by the end of the round, while during 322 of the 352 rounds (88.0%) no critical items remained unchecked. The average percentage of completed checklist items and unnoticed critical items per day were 97.4% (range 88.8 – 100.0) and 2.0% (range 0.0 – 9.8), respectively.

Qualitative outcomes

User experience

The AttrakDiff questionnaire was completed by 21 participants after each period. The TAM-2 based questionnaire was completed by 18 participants after the control period and 21 participants after the intervention period. Participants' characteristics for the questionnaires were similar in both groups (eTable 5-6, Supplement 1).

The ease-of-use factors of the DCC were rated higher compared to the paper checklist, while the pragmatic factors remained similar (*Figure 3* and eTable 5, Supplement 1). Participants valued the DCC as more innovative and inventive. The DCC was the preferred tool of the participants to accomplish their goals during the rounds. No differences were found for most domains of TAM-2. Only the median score for "facilitation" was higher with the DCC compared to the paper checklist (eTable 6, Supplement 1).

Semi-structured interviews

Independent researchers coded quotations from nine interviews (word count range: 1845 – 4646 words) with an average of 64 coded quotations per interview (range: 47-97 quotations), resulting in a total of 577 quotations. A wide variety of opinions were expressed by the interviewed clinicians regarding their attitudes towards checklists, the DCC and its job relevance, perceived ease of use and perceived usefulness, and more generally towards technology application in healthcare. *Table 3* provides the most illustrative quotes under each category of the TAM-2.

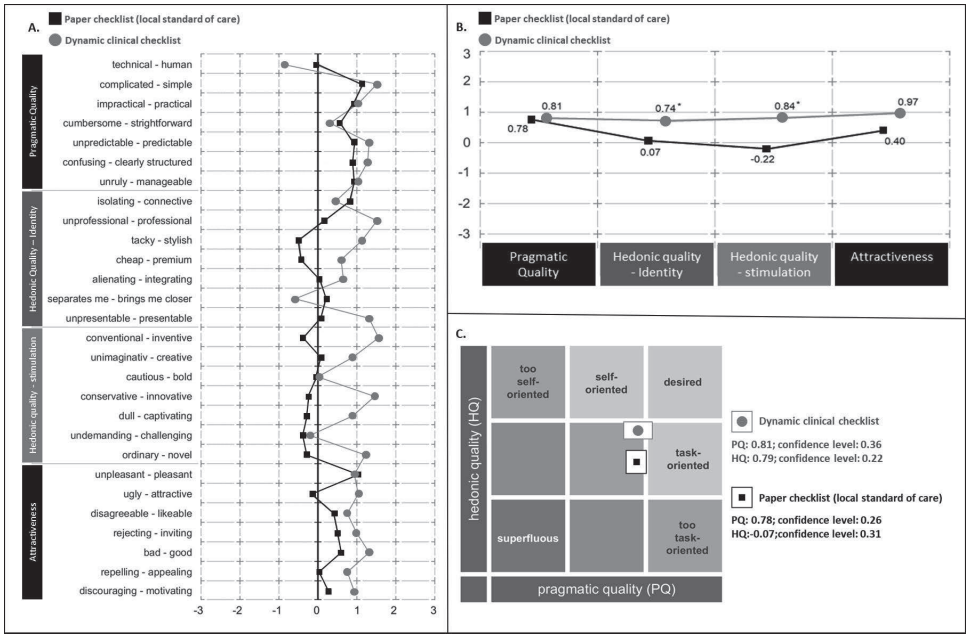


Figure 3. Results of the AttrakDiff questionnaire.
Results of the AttrakDiff questionnaire: (A) mean scores for each element of the questionnaire, (B) mean score for each domain, and (C) and the overall score of the ease-of-use (hedonic) and pragmatic qualities.
Abbreviations: PQ = Pragmatic Quality; HQ-I = Hedonic Quality – Identity; HQ-S = Hedonic quality - stimulation; ATT = Attractiveness.

Most clinicians found the DCC easy to use. They did not receive training on how to use it, but they expressed that the DCC was intuitive to use. Furthermore, they found the progress circle a clear and strong element of the DCC which motivated them to complete the checklist (Quote 1). The most frequent users also expressed some negative comments, like the system being too slow or requesting to repeat logging in (Quote 2).

In terms of job relevance, most clinicians agree that the content of the DCC was relevant for their work (Quote 3). Two clinicians stated that the DCC contains too many irrelevant topics, making it a time consuming effort. Five clinicians appreciated the support provided by the DCC for anti-coagulation for which the clinical decision making depends on a comprehensive guideline that regularly changes (Quote 4).

Most clinicians perceived the DCC very useful since they believed it could prevent mistakes (Quote 5). The majority of the clinicians expressed that the DCC improves the adherence to the checklist. However, they were not yet convinced that this improvement

Table 3 – Quotes of interviewed participants after the intervention period

TAM-2 categories	Quotes		Participant
Perceived ease of use	Q1	*"The layout is good. The DCC is very clear with that little round circle. I find that a positive thing. It is quick. It is not a very slow system. […] I think those are the real benefits."*	Female, intensivist, <5 years of experience
	Q2	*"…I found it annoying that it sometimes jumps out, logged out and that I have to log back in…Or if I accidentally press on it, it switches off."*	Male, intensivist, <5 years of experience
Job relevance	Q3	*"I don't want to be dependent on such a machine when thinking about a patient. But I am a human being and I make mistakes. I slip up and then a safety net is welcome."*	Female, intensivist, 5-10 years of experience
	Q4	*"I think that the greatest added value is related to medication. That's where I found the extra information [from the DCC] always valuable."*	Female, physician assistant, <5 years of experience
Perceived usefulness	Q5	*"It's a sort of check for myself, if I am not missing anything, if I have thought about everything. It provides structure. Yes, it is an aid. I feel more reassured when using the checklist."*	Female, physician assistant, <5 years of experience
	Q6	*"To improve the health of patients is quite a stretch. I think it increases the odds of people following protocol. But if that improves the odds of patients improving? I don't know. Maybe sometimes yes, sometimes no. But protocols are there for a reason and it is important to follow them"*	Male, physician assistant, 5-10 years of experience
Beliefs and attitudes	Q7	*"The danger is if you only are trained with cognitive support tools and in the end, no one can really help you with that. Identifying the main topics and details, and being able to see the endpoint, that is where we need to focus."*	Male, intensivist, <5 years of experience
	Q8	*"Decision support is valuable. It is more valuable than checklists. The DCC is now too much of a checklist."*	Male, intensivist, 5-10 years of experience

Abbreviations: DCC = dynamic clinical checklists

would translate into improved patient outcomes because the DCC's content mainly covers protocolized care processes but does not support the clinical decision making concerning the treatment of the underlying life-threatening diseases (Quote 6).

A few clinicians expressed some fear and resistance to more technology in health care as it could make clinicians too dependent on technology (Quote 7). They appreciate the structure that the DCC provides for education, but fear that it can prevent doctors from thinking independently. On the contrary, others foresee a bright future for the DCC and more advanced clinical decision support tools. Clinicians that were regular paper checklist users expressed the most positive opinions towards the DCC and were most likely to use it. Overall, the DCC was perceived as a suitable tool to replace paper checklists, however, its usability can be improved (Quote 8).

DISCUSSION

In this prospective before-and-after mixed-method study we observed that compliance with the best eligible practice improved after a DCC was implemented during ICU ward rounds. Use of the DCC was associated with a significant reduction in ICU length of stay and fewer days with prescribed empiric antibiotics. In the intervention period the increased number of days with intravenous continuous sedatives translated into less unacceptable levels of critical care pain scores (CPOT>2). Overall, physicians valued the DCC as an attractive and innovative technology. Although in questionnaires the DCC's effect on usability was rated similar to the paper checklist, the majority of the physicians mentioned that the DCC was easily applied in daily practice and has more future potential compared to paper checklists.

The compliance rate of the paper checklist items in the control period was similar to other studies and the previous simulation pilot study.[9, 11, 28-30] As compared to the pilot study, similar high rates of checked items were found in this study, but now for an intervention period of eight weeks in real practice.[11] These high rates provide reassurance that physicians have considered the presented items and chose to follow or intentionally deviate from protocol in the interest of their individual patient. The observed improvement is in line with results of most other healthcare studies evaluating electronic checklists.[10, 11, 31] However, these studies compared electronic checklists with no checklist or were simulation studies. Thus it remains difficult to determine if specific features of electronic checklists were responsible for the higher compliance rates besides the general impact of having a checklist as a memory aid.[10] Our present study suggests that the observed higher compliance rate is the result of the specific features

of an electronic checklist with dynamic properties since it compared the DCC with a paper checklist in real practice. The improved rate can partly be explained by the DCC's ability to automatically check items. Approximately a third of the checkable items was checked automatically resulting in shorter checklists with a more relevant content.

The present study showed that the use of the DCC was associated with a reduced length of stay in the ICU and several improved care processes. However, some care processes included in the DCC did not result in improvements. Nor could we reproduce the reduction of pharmacists' phone calls due to violated pharmacological clinical rules found in the previous simulation study.[11] This discrepancy might be explained by the different study designs. In the simulation study some flaws were deliberately implemented and occurred therefore in every simulation. In real practice, however, there are more occasions over the day and night that the violation of a pharmacological clinical rules can be noticed and corrected. This might explain why fewer violations were found during the ward rounds and this sample size was probably too small to detect a significant difference, though the absolute number of violations tended to be smaller in the intervention period. The lack of improvement in the other care processes is in line with the findings of other quality improvement multifaceted approach studies on the ICU.[9, 28 32-34] Similar to these kind of studies, our findings need further study since several potential factors may be involved:(1) the before-and-after design of this study might have been influenced by secular trends,[35](2) the inclusion of a broader patient population than the studies that found an effect of a specific care process,(3) even though some care processes are recommended by ICU guidelines, their effect on patient outcomes are still undetermined,[36, 37](4) the study was not powered for particular secondary outcomes (e.g. a smaller number of violated pharmacological clinical rules was found in this real practice study compared to the previous simulation study),[11](5) both the period of intervention and follow-up time were too short to find measurable effects.[38]

The discrepancy of improved ease-of-use with no effect on usability, as observed with the questionnaires, was unexpected. Insufficient training and time for introduction, or a complicated user-interface seem not be causing this difference since most users perceived the DCC as intuitive except for some hitches. In addition, in the interviews users appreciated that the DCC provided useful suggestions and tried to prevent mistakes that they thought would have otherwise remained unnoticed. This opinion of a checklist's purpose is in line with a previous qualitative study evaluating a paper checklist for ICU ward rounds.[39] In the present study users acknowledged errors could be prevented, especially for items of care processes that are based on complicated and frequently updated guidelines, like the anticoagulation-related management guideline. On the contrary, users indicated that the DCC still functioned too much as

a checklist for care processes instead of a cognitive aid supporting decision making at the bedside. In their opinion the dynamic properties of this DCC have the potential to fulfil this expectation, although some also argued that clinicians should not become too dependent on technology.

Future development of the DCC and studies should focus on:(1) the capability of retrieving more data and trends from multiple sources as input for algorithms, including wearable devices,(2) implementing and validating multiple interacting DCCs for various clinicians within patients' clinical pathways,(3) the use of more sophisticated algorithms based on machine learning for which the DCC can also retrieve reliable relevant missing data,(4) the DCC's capability of improving the translation and application of the available or new medical knowledge and evidence into clinical practice, such as during emerging pandemics, and (5) improving user experience.

This study has several limitations. The single centre nature of the study decreases its external validity. Most biases are inherent to both the before-and-after and mixed-method design of this study. Although the baseline characteristics seemed balanced, subtle differences might have introduced selection bias. Regression to the mean might have occurred since multiple rounds were observed in patients admitted for more than one day. We tried to reduce the Hawthorne effect through discrete observation of participant behaviour. Measurement bias could have occurred in the control period since the output of the DCC was used to judge if items needed to be discussed. Participation in the questionnaires and interviews was voluntary and participants might have tried to please the investigators.

CONCLUSION

In an ICU with high baseline standard of care the introduction of a DCC for ICU ward rounds improved the compliance to best eligible practice and was associated with a reduction in ICU length of stay, daily use of antibiotics, and pain observation scores. The DCC needs further refinement in terms of usability and its dynamic properties to fulfil physicians' expectations of a patient-centred and user-specific cognitive aid.

REFERENCES

1. Haynes AB, Weiser TG, Berry WR, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *The New England journal of medicine* 2009; **360**: 491-9
2. Bergs J, Lambrechts F, Simons P, et al. Barriers and facilitators related to the implementation of surgical safety checklists: a systematic review of the qualitative evidence. *BMJ quality & safety* 2015; **24**: 776-86
3. de Vries EN, Prins HA, Crolla RM, et al. Effect of a comprehensive surgical safety system on patient outcomes. *The New England journal of medicine* 2010; **363**: 1928-37
4. Haugen AS, Softeland E, Almeland SK, et al. Effect of the World Health Organization checklist on patient outcomes: a stepped wedge cluster randomized controlled trial. *Annals of surgery* 2015; **261**: 821-8
5. Bergs J, Hellings J, Cleemput I, et al. Systematic review and meta-analysis of the effect of the World Health Organization surgical safety checklist on postoperative complications. *The British journal of surgery* 2014; **101**: 150-8
6. Fourcade A, Blache JL, Grenier C, Bourgain JL, Minvielle E. Barriers to staff adoption of a surgical safety checklist. *BMJ quality & safety* 2012; **21**: 191-7
7. van Klei WA, Hoff RG, van Aarnhem EE, et al. Effects of the introduction of the WHO "Surgical Safety Checklist" on in-hospital mortality: a cohort study. *Annals of surgery* 2012; **255**: 44-9
8. Borchard A, Schwappach DL, Barbir A, Bezzola P. A systematic review of the effectiveness, compliance, and critical factors for implementation of safety checklists in surgery. *Annals of surgery* 2012; **256**: 925-33
9. Cavalcanti AB, Bozza FA, Machado FR, et al. Effect of a Quality Improvement Intervention With Daily Round Checklists, Goal Setting, and Clinician Prompting on Mortality of Critically Ill Patients: A Randomized Clinical Trial. *Jama* 2016; **315**: 1480-90
10. Kramer HS, Drews FA. Checking the lists: A systematic review of electronic checklist use in health care. *Journal of biomedical informatics* 2017; **71S**: S6-S12
11. De Bie AJR, Nan S, Vermeulen LRE, et al. Intelligent dynamic clinical checklists improved checklist compliance in the intensive care unit. *British journal of anaesthesia* 2017; **119**: 231-8
12. Webster CS. Checklists, cognitive aids, and the future of patient safety. *British journal of anaesthesia* 2017; **119**: 178-81
13. Nan S, De Bie A, Zhang S, Korsten H, Lu X, Duan H. Identify Facilitators and Challenges in Computerized Checklist Implementation. *Studies in health technology and informatics* 2019; **264**: 1737-8
14. Nan S, Van Gorp P, Lu X, et al. A meta-model for computer executable dynamic clinical safety checklists. *BMC medical informatics and decision making* 2017; **17**: 170
15. Nan S VGP, Korsten H, Vdovják R, Kaymak U. Tracebook: a dynamic checklist support system. *IEEE 27th International Symposium on Computer-Based Medical Systems* 2014: 48-51
16. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007; **370**: 1453-7

17. Kreeftenberg HG, Aarts JT, de Bie A, Bindels A, Roos AN, van der Voort PHJ. An alternative ICU staffing model: implementation of the non-physician provider. *The Netherlands journal of medicine* 2018; **76**: 176-83
18. Vincent JL. Give your patient a fast hug (at least) once a day. *Critical care medicine* 2005; **33**: 1225-9
19. Papadimos TJ, Hensley SJ, Duggan JM, et al. Implementation of the "FASTHUG" concept decreases the incidence of ventilator-associated pneumonia in a surgical intensive care unit. *Patient safety in surgery* 2008; **2**: 3
20. de Clercq PA, Hasman A, Blom JA, Korsten HH. Design and implementation of a framework to support the development of clinical guidelines. *International journal of medical informatics* 2001; **64**: 285-318
21. Scheepers-Hoeks AMJW GR, Neef C, Doppen AMJ, Korsten EHM. Preventive Prescribing of Laxatives for Opioid-induced Constipation Using Electronic Clinical Rule Implementation by Cinical Pharmacists. *Advances in Pharmacoepidemiology & Drug Safety* 2014; **3**
22. van der Linden CM, Jansen PA, van Marum RJ, Grouls RJ, Egberts TC, Korsten EH. An electronic system to document reasons for medication discontinuation and to flag unwanted represcriptions in geriatric patients. *Drugs & aging* 2012; **29**: 957-62
23. Barr J, Fraser GL, Puntillo K, et al. Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit. *Critical care medicine* 2013; **41**: 263-306
24. Ely EW, Truman B, Shintani A, et al. Monitoring sedation status over time in ICU patients: reliability and validity of the Richmond Agitation-Sedation Scale (RASS). *Jama* 2003; **289**: 2983-91
25. Hassenzahl M, Platz A, Burmester M, Lehner K. Hedonic and ergonomic quality aspect determine a software's appeal. 2000
26. Venkatesh V, Davis FD. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science* 2000; **46**: 186-204
27. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 1995; **57**: 289-300
28. Carlos WG, Patel DG, Vannostrand KM, Gupta S, Cucci AR, Bosslet GT. Intensive care unit rounding checklist implementation. Effect of accountability measures on physician compliance. *Annals of the American Thoracic Society* 2015; **12**: 533-8
29. de Vries EN, Prins HA, Bennink MC, et al. Nature and timing of incidents intercepted by the SURPASS checklist in surgical patients. *BMJ quality & safety* 2012; **21**: 503-8
30. Pucher PH, Aggarwal R, Almond MH, Darzi A. Surgical care checklists to optimize patient care following postoperative complications. *American journal of surgery* 2015; **210**: 517-25
31. Sevilla-Berrios R, O'Horo JC, Schmickl CN, et al. Prompting with electronic checklist improves clinician performance in medical emergencies: a high-fidelity simulation study. *International journal of emergency medicine* 2018; **11**: 26

32. Byrnes MC, Schuerer DJ, Schallom ME, et al. Implementation of a mandatory checklist of protocols and objectives improves compliance with a wide range of evidence-based intensive care unit practices. *Critical care medicine* 2009; **37**: 2775-81
33. Weiss CH, Moazed F, McEvoy CA, et al. Prompting physicians to address a daily checklist and process of care and clinical outcomes: a single-site study. *American journal of respiratory and critical care medicine* 2011; **184**: 680-6
34. Scales DC, Dainty K, Hales B, et al. A multifaceted intervention for quality improvement in a network of intensive care units: a cluster randomized trial. *Jama* 2011; **305**: 363-72
35. Ko HC, Turner TJ, Finnigan MA. Systematic review of safety checklists for use by medical care teams in acute hospital settings--limited evidence of effectiveness. *BMC health services research* 2011; **11**: 211
36. Aitken LM, Bucknall T, Kent B, Mitchell M, Burmeister E, Keogh SJ. Protocol-directed sedation versus non-protocol-directed sedation in mechanically ventilated intensive care adults and children. *The Cochrane database of systematic reviews* 2018; **11**: CD009771
37. Wang L, Li X, Yang Z, et al. Semi-recumbent position versus supine position for the prevention of ventilator-associated pneumonia in adults requiring mechanical ventilation. *The Cochrane database of systematic reviews* 2016: CD009946
38. van der Voort PHJ. Quality improvement in the ICU: treat first what kills first. *Journal of thoracic disease* 2017; **9**: E310-E2
39. Hallam BD, Kuza CC, Rak K, et al. Perceptions of rounding checklists in the intensive care unit: a qualitative study. *BMJ quality & safety* 2018; **27**: 836-43

SUPPLEMENTAL DIGITAL CONTENT 1

eTable 1. 2018 characteristics of the Intensive Care Unit of the Catharina Hospital in Eindhoven, the Netherlands.

ICU Characteristics	Number (%)
ICU beds (maximum)	37
Fulltime-equivalent:	
Intensivists (maximum)	22
ICU nurses (maximum)	152
ICU admissions:	**2995**
Medical	871 (29.1)
Surgical	
Acute surgery	288 (9.6)
Elective surgery	1833 (61.2)
Other type of admission	3 (0.1)
Surgical admissions:	**2121**
Cardiac surgery	1209 (57.0)
Aneurysm surgery	65 (3.1)
Thoracotomy	101 (4.8)
Gastro-intestinal malignancy	212 (10.0)
Other surgery	534 (25.2)
Estimated risk of mortality at admission based on APACHE IV):	**2995**
Low risk (<30%)	2574 (85.9)
Moderate risk (30 - <70%)	171 (5.7)
High risk (≥70%)	71 (2.4)
Unknown	179 (6.0)
Standardized Mortality Ratio based on the APACHE IV model:	
Overall	0.76
Medical admissions:	0.81
Community acquired pneumonia	1.04
Sepsis	0.58
Out of hospital cardiac arrest	1.50
Surgical admission:	
Acute surgery	0.80
Elective surgery	0.61

ICU Characteristics	Number (%)
APACHE IV score:	
Low risk (<30%)	0.64
Moderate risk (30 - <70%)	0.73
High risk (≥70%)	0.98
Readmission rate (corrected and based on Dutch average)	
Overall	0.7
Within 48hours after discharge	0.4
ICU Length of stay (days)	
Overall	
Medical admissions:	1.0
Community acquired pneumonia	1.8
Sepsis	2.0
Surgical admission:	
Acute surgery	1.3
Elective surgery	1.0
APACHE IV score:	
Low risk (<30%)	1.0
Moderate risk (30 - <70%)	2.8
High risk (≥70%)	4.1
Duration of mechanical ventilation (days)	
Overall	0.2
Medical admissions:	1.7
Community acquired pneumonia	4.4
Sepsis	0.5
Surgical admission:	
Acute surgery	0.4
Elective surgery	0.2
APACHE IV score:	
Low risk (<30%)	0.2
Moderate risk (30 - <70%)	2.1
High risk (≥70%)	1.6

eTable 2. List of items to check during observations of the ICU ward rounds.

Patient name	ID	ICU Room	Date:	
Items	**Reported in the EMR**	**Discussed at the bedside**	**Advice of TraceBook considered (yes/no)**	**Note**
Feeding (not applicable during first 48 hours with a normal diet prescribed)				N.A?
Sufficient calories				N.A?
Defecation				N.A?
Pain and use of analgesics (not applicable if VAS <4 in the first 48hrs after surgery)				N.A? VAS: CPOT:
Use of sedatives				N.A? RASS:
Delirium (only applicable if the CAM-ICU was positive <24hours)/ RASS>0 / scored complication or prescribed medication)				N.A? CAM-ICU:
Thrombosis prophylaxis				
Head lift >30^{o}				N.A?
Stress ulcer prophylaxis				N.A?
Glucose in range				N.A?
SDD applied conform protocol				N.A?
Antibiotics prescribed or required; positive cultures discussed.				N.A?
Need for invasive catheters				
Decubitus (Only applicable if registered)				N.A?
Medical history discussed				

Patient name	ID	ICU Room	Date:	
Items	**Reported in the EMR**	**Discussed at the bedside**	**Advice of TraceBook considered (yes/no)**	**Note**
Estimated blood results discussed				
Performed imaging discussed				
Pharmacological clinical rules applicable and discussed?		**Applicable (which one):** 1. 2. 3. 4. 5.	**Discussed:** 1. 2. 3. 4. 5.	
Complications of the last 24 hours discussed and scored?	Scored in EMR:	Discussed:		
Prescribed medication reviewed				
Conclusion discussed and reported in EMR				
Plan of the day discussed and reported in EMR				
Note				

Abbreviations: ICU=intensive care unit; EMR=electronic medical record; CAM-ICU=Confusion Assessment Method – ICU; RASS= Richmond Agitation-Sedation Scale.

eTable 3. Baseline characteristics of the included patients

	Control period (*n* = 197)	**Intervention period (*n* = 211)**
Age, years	69 (58 – 76)	68 (59 – 74)
Male, sex	133 (67.5)	146 (69.2)
BMI, kg m^{-2}	26.6 (23.9 – 29.8)	26.1 (23.8 – 28.9)
Quick SOFA at admission	1 (1 – 2)	1 (1 – 2)
SAPS 2	35 (28 – 47)	34 (27 – 42)
APACHE II	17 (13 – 22)	16 (13 – 20)
APACHE IV	45 (34 – 59)	43 (34 – 58)
EuroSCORE II*	5.0 (3.0 – 7.5)	5.0 (3.0 – 7.0)
Charlson Comorbidity Score	4 (3 – 6)	4 (2 – 5)
Diabetes Mellitus	44 (22)	36 (17)
Chronic kidney disease (creatinine >3 mg dL^{-1} (0.27 mmol L^{-1}))	23 (12)	16 (8)
ECMO	2 (1)	4 (2)
Mechanically ventilation required	142 (72.1)	149 (71.0)
Vasopressor need first 24 hours	118 (59.9)	140 (66.4)
Inotropic need first 24 hours	36 (18.3)	44 (20.9)
Admission source		
Emergency department	51 (26)	43 (20)
Ward	27 (14)	23 (11)
Postsurgical	119 (60)	145 (69)
Type of surgery		
Elective	110 (55.8)	121 (57.3)
Acute	23 (11.7)	26 (12.3)
Reason for admission		
Postsurgical	126 (64)	139 (66)
Shock		
Septic	7 (4)	8 (4)
Hypovolemic	3 (1)	10 (5)
Cardiogenic	4 (2)	4 (2)
Distributive	0 (0)	1 (0)
Respiratory distress	19 (10)	17 (8)
Non-invasive ventilation	1 (0)	0 (0)
Hemodynamic monitoring otherwise	10 (5)	15 (7)
Sepsis	11 (6)	5 (2)
Cardiac arrest	5 (3)	6 (3)
Telemetry	3 (1)	6 (3)
Intoxication	5 (3)	0 (0)
Thrombolysis	3 (1)	0 (0)

	Control period (*n* = 197)	**Intervention period (*n* = 211)**
Admission specialty		
Cardiothoracic surgery	88 (44)	104 (49)
General surgery	48 (24)	43 (20)
Internal medicine	18 (9)	15 (7)
Cardiology	10 (5)	18 (9)
Pulmonology	11 (6)	8 (4)
Gastroenterology	4 (2)	13 (6)
Gynaecology and Urology	8 (4)	4 (2)
Neurology	5 (3)	6 (3)
Others	5 (3)	0 (0)
Glasgow coma scale ≤8 at admission	8 (5)	12 (6)
PaO_2 at admission	90 (76 - 119)	92 (73 - 114)
pH at admission	7.36 (7.32 – 7.41)	7.36 (7.29 – 7.41)
Highest lactate level in the first 24 hours	2.3 (1.4 – 5.5)	2.3 (1.4 – 3.5)
Lowest MAP in the first 24 hours (mm Hg)	55 (47 – 63)	53 (50 – 63)
Heart rate at admission (beats per minute)	79 (68 – 96)	80 (70 – 90)
Highest Respiratory rate the first 24 hours (breaths per minute)	23 (19 – 26)	22 (19 -26)

Data are median (25 – 75% quartile) or No (%)

PBW: predicted body weight; BMI: body mass index; SAPS: Simplified Acute Physiology Score; APACHE: Acute Physiology and Chronic Health Evaluation; COPD: chronic obstructive pulmonary disease; OSA: obstructive sleep apnea; CVD: cerebrovascular disease; TIA: transient ischemic attack; NYHA: New York Heart Association; LVEF: left-ventricular ejection fraction; CABG: coronary artery bypass graft; CK-MB: Creatine Kinase MB Isoenzyme.

** Patients with elective cardiac surgery (n=77 in control and n=81 in the intervention period)*

eTable 4. Post Hoc multiple comparisons test: False Discovery Rate correction

Comparison (control vs intervention period)	Raw p-value	Adjusted p-value	Test number (i)	FDR cut-offs*	Significant
Percentage of checked items	< 0.001	0.030	1	0.002	Yes
Percentage of unchecked critical items	< 0.001	0.015	2	0.003	Yes
Intravenous sedatives^	< 0.001	0.010	3	0.005	Yes
Antibiotics^^	< 0.001	0.008	4	0.007	Yes
CPOT (0-6)	< 0.001	0.006	5	0.008	Yes
CPOT >2	0.001	0.005	6	0.010	Yes
Length of stay: ICU	0.011	0.047	7	0.012	Yes
Length of stay: hospital	0.014	0.053	8	0.013	Yes
Opiates; days prescribed on regular base	0.05	0.167	9	0.015	No
CRBSI	0.07	0.210	10	0.017	No
Hyperglycemia	0.07	0.191	11	0.018	No
Days without defecation for at least >48hours, median (min – max)	0.11	0.275	12	0.020	No
Number of alerts (Medication related rules)	0.13	0.300	13	0.022	No
Paracetamol; days prescribed on regular base	0.19	0.407	14	0.023	No
RASS	0.22	0.440	15	0.025	No
Gastro-intestinal bleedings	0.25	0.469	16	0.027	No
VAS	0.37	0.653	17	0.028	No
Hospital acquired pneumonia	0.49	0.817	18	0.030	No
Registered complications	0.74	1.168	19	0.032	No
No PPI, while indicated	0.77	1.155	20	0.033	No
Mortality: 90 day	0.74	1.057	21	0.035	No
VAS >4	0,8	1.091	22	0.037	No
Number of interventions based alerts	0.83	1.083	23	0.038	No
Invasive ventilation time, hours (patients with >24 hours of ventilation)	0.84	1.050	24	0.040	No
Hypoglycaemia	0.85	1.020	25	0.042	No
Phone calls of pharmacist to ICU clinician	0.88	1.015	26	0.043	No
Number of relevant alerts (Medication related rules)^^^	0.88	0.978	27	0.045	No
Mortality: 30 day	0.90	0.964	28	0.047	No
Mortality: ICU	0.98	1.014	29	0.048	No
Invasive ventilation time, hours (overall group)	0.98	0.980	30	0.050	No

Abbreviations: FDR = False Discovery Rate; CRBSI = central-venous-catheter-related bloodstream infections; PPI = proton pump inhibitor

** Adjusted p-value = p-value * (m/i); m=26. ** FDR cut-off = α (=0.05) * i/m); m=26. ^ Propofol or Midazolam. ^^ Days with only selective digestive decontamination excluded. ^^^ Determined by the hospital pharmacist on duty*

eTable 5. Participants' characteristics and results of the Attrakdiff questionnaire.
Participants' characteristics and comparison of the median scores for each domain of the AttrakDiff questionnaire, with Cronbach's alpha, between the paper checklist (control period) and the dynamic clinical checklist (intervention period).

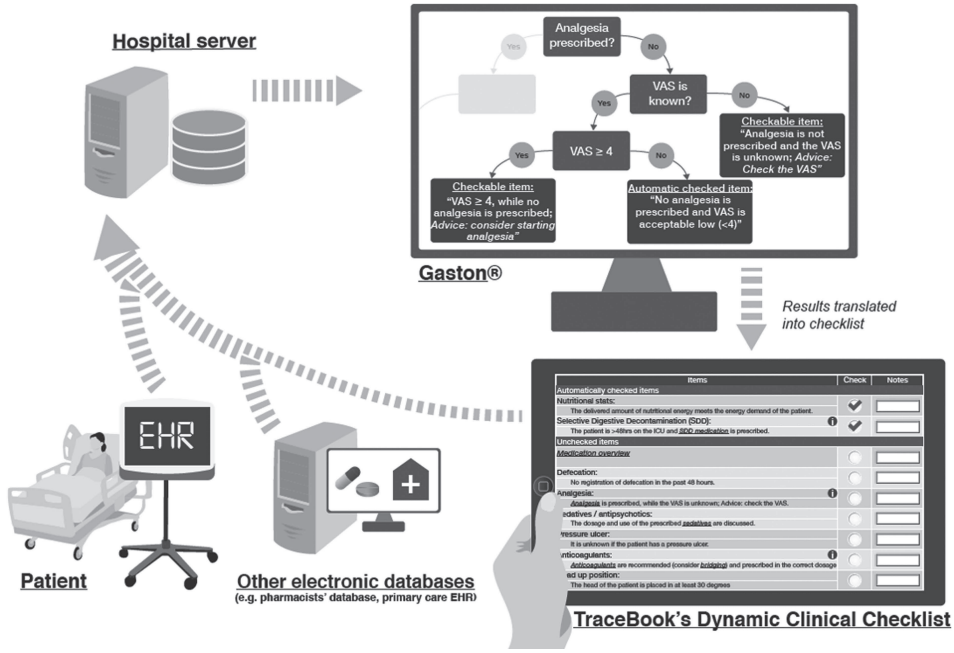
AttrakDiff questionnaire 7 items per domain; Likert scale: -3–3	**Control period (n=21)**	**Intervention period (n=21)**	**p-value**	**95% CI**
Response rate (%)	58	70		
Male, n (%)	11 (52)	12 (57)	1.00	
Age <40 years, n (%)	11 (52	10 (48)	1.00	
Job, n (%)				
Intensivist	10 (48)	9 (43)		
Resident	5 (24)	6 (29)		
ICU physician assistant	6 (29)	6 (29)	0.93	
ICU Experience, yrs	5.5 (1.8 – 15.5)	9.0 (2.5 – 15.0)	0.99	
Pragmatic quality (PQ)				
Cronbach's alpha	0.53	0.84		
Mean (SD)	0.78 (0.61)	0.81 (0.85)	0.882	-0.43 – 0.50
Hedonic Quality – identity (HQ-I)				
Cronbach's alpha	0.69	0.68		
Mean (SD)	0.08 (0.82)	0.74 (0.61)	0.005	0.21 – 1.12
Hedonic Quality – stimulation (HQ-S)				
Cronbach's alpha	0.80	0.67		
Mean (SD)	-0.22 (0.74)	0.84 (0.56)	<0.001	0.65 – 1.47
Attractiveness (ATT)				
Cronbach's alpha	0.89	0.94		
Mean (SD)	0.40 (0.83)	0.97 (1.01)	0.052	-0.01 – 1.15

eTable 6. Participants' characteristics and results of the TAM-2 based questionnaire.
Participants' characteristics and comparison of the median scores for each domain of the TAM-2 based questionnaire between the paper checklist (control period) and the DCC (intervention period).

	Control period (n=18)	Intervention period (n=21)	z-score	p-value
Response rate (%)	50	70		-
Male, n (%)	11 (52%)	14 (61%)		0.79
Age, years	40 (31–47)	40 (36–48)		0.60
Job				0.91
Intensivist	9 (45%	12 (52%)		
Resident	5 (25%)	6 (26%)		
ICU physician assistant	6 (30%)	5 (22%)		
Experience, years	4.8 (0.5 – 11.8)	5.0 (1.0 – 13.5)		0.49
Subjective Norm				
Median (1st & 3th quartile)	4 (3 – 4.25)	4 (3 – 4)	-0.99	0.32
Imaging				
Median (1st & 3th quartile)	3.5 (3.29 – 3.63)	3.38 (3.23 – 3.65)	-0.42	0.67
Job Relevance				
Median (1st & 3th quartile)	4 (3.69 – 4.06)	4 (3.5 – 4)	-0.97	0.33
Output Quality				
Median (1st & 3th quartile)	3.43 (3.25 – 3.61)	3.43 (3.21 – 3.71)	-0.04	0.97
Results Demonstrability				
Median (1st & 3th quartile)	3.80 (3.60 – 4.00)	3.8 (3.40 – 4.00)	-0.72	0.47
Perceived ease of use				
Median (1st & 3th quartile)	3.63 (3.44 – 3.92)	3.83 (3.5 – 4.00)	-0.87	0.39
Facilitation				
Median (1st & 3th quartile)	3.00 (2.50 – 3.50)	4.00 (3.25 – 4.00)	-2.50	0.01
Perceived usefulness				
Median (1st & 3th quartile)	153.50 (144.00 – 160.50)	152.00 (146.00 – 159.00)	-0.24	0.81
Acceptance				
Median (1st & 3th quartile)	159.5 (149.75 – 167.25)	160 (154.5 – 165.5)	-0.03	0.98
Grade (1-5)				
Median (1st & 3th quartile)	3.00 (3.00 – 4.00)	4.00 (3.00 – 4.00)	-0.98	0.33

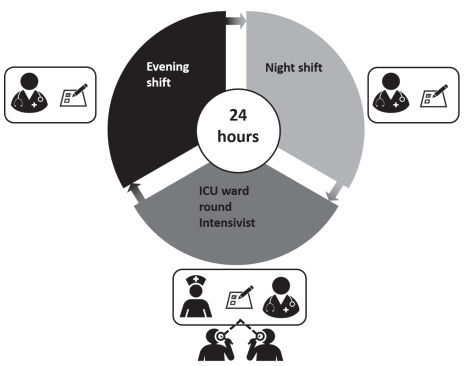
eTable 7: Topic guide for the qualitative interviews.
Topic guide for the qualitative semi-structured interviews in the two weeks after the intervention period.

Topic		Specific elements
1	Introduction	Study goals and procedures Privacy and confidentiality Pinging informed consent
2	Background variables	Specialism, years of experiences Use of digital technology in daily life
3	Checklist usage	Usage of regular paper checklist (FAST HUG) Motivation for study participation
4	Expectations	Expectations of the Dynamic Clinical Checklist (DCC) before the study
5	Usage of DCC	Usage of DCC during intervention period Moments of (non) usage Barriers for usage
6	Perceived usefulness	Relevance for job Advantages and disadvantages Expected outcomes Most appreciated elements Least appreciated elements
7	Perceived usability	Ease of understanding Need for support Attractiveness Speed
8	Technology in health care	
9	Closing; thank you; end	

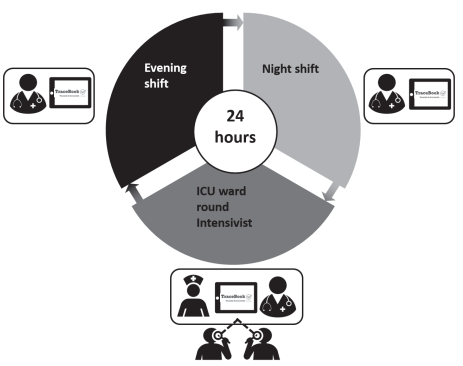


eFigure 1. A schematic overview of how a TraceBooks DCC is composed.

The information about a patient is gathered by GASTON on the hospital server. The rule engines of GASTON and TraceBook decide, based on algorithms, which rule is relevant and can become a checkable item or automatically checked item. A small part of the algorithm of prescribing analgesia based on the pain rating scale is shown, which can provide an automatically checked item if no analgesia is prescribed and the pain rating score is low (<4). DCC, dynamic clinical checklist; e.g., example; EHR, electronic health record; VAS, visual analogue scale.



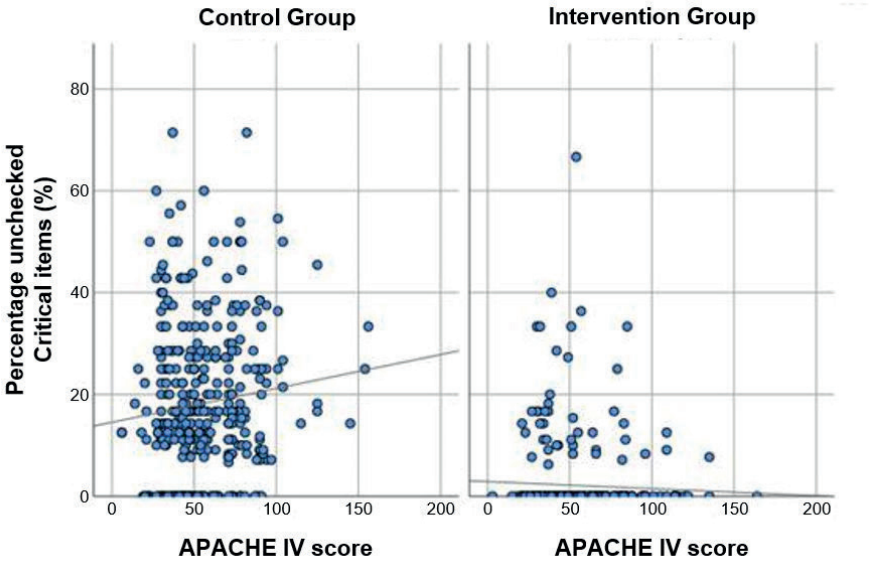
A. Standard operating procedure during control period; TraceBook is only available for the observers.



B. Standard operating procedure during intervention period; TraceBook is available for ICU clinicians and the observers.

eFigure 2. Standard operating procedures

Standard operating procedures during the control (A) and intervention (B) periods.



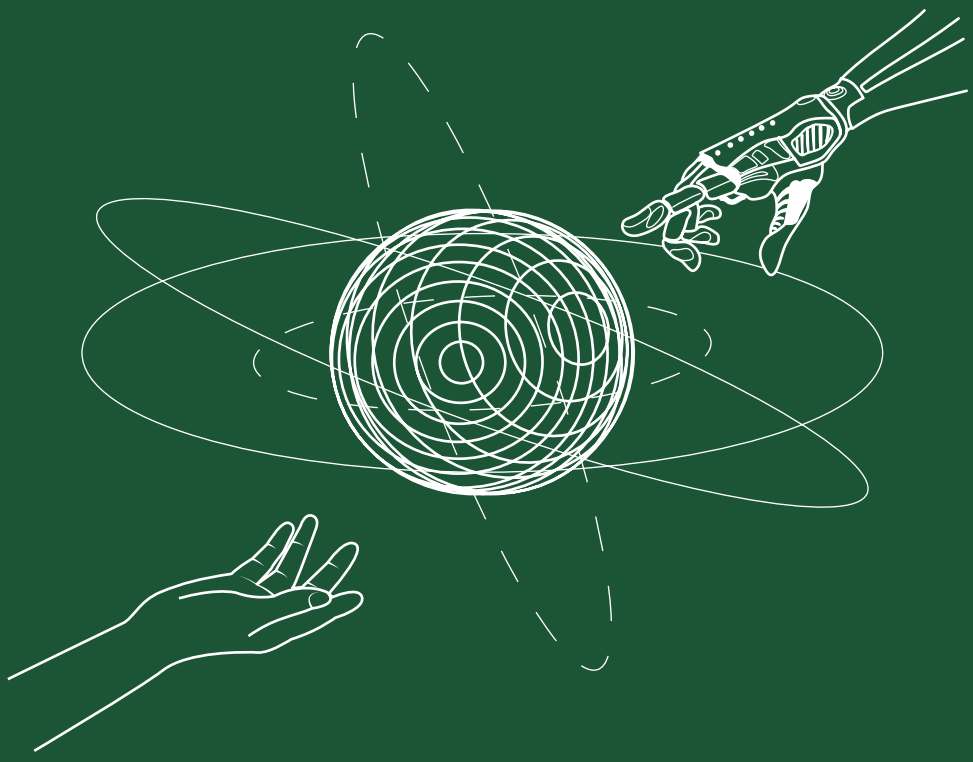
eFigure 3. Correlation plots of the APACHE IV scores in the control and the intervention period.

APPENDIX 1. DISTRIBUTION OF NORMALITY – TESTS OF NORMALITY AND Q-Q PLOTS

For a detailed description on the tests of normality please refer to the supplementary data on the website of the publishing journal.

SUPPLEMENTAL 2: STUDY PROTOCOL

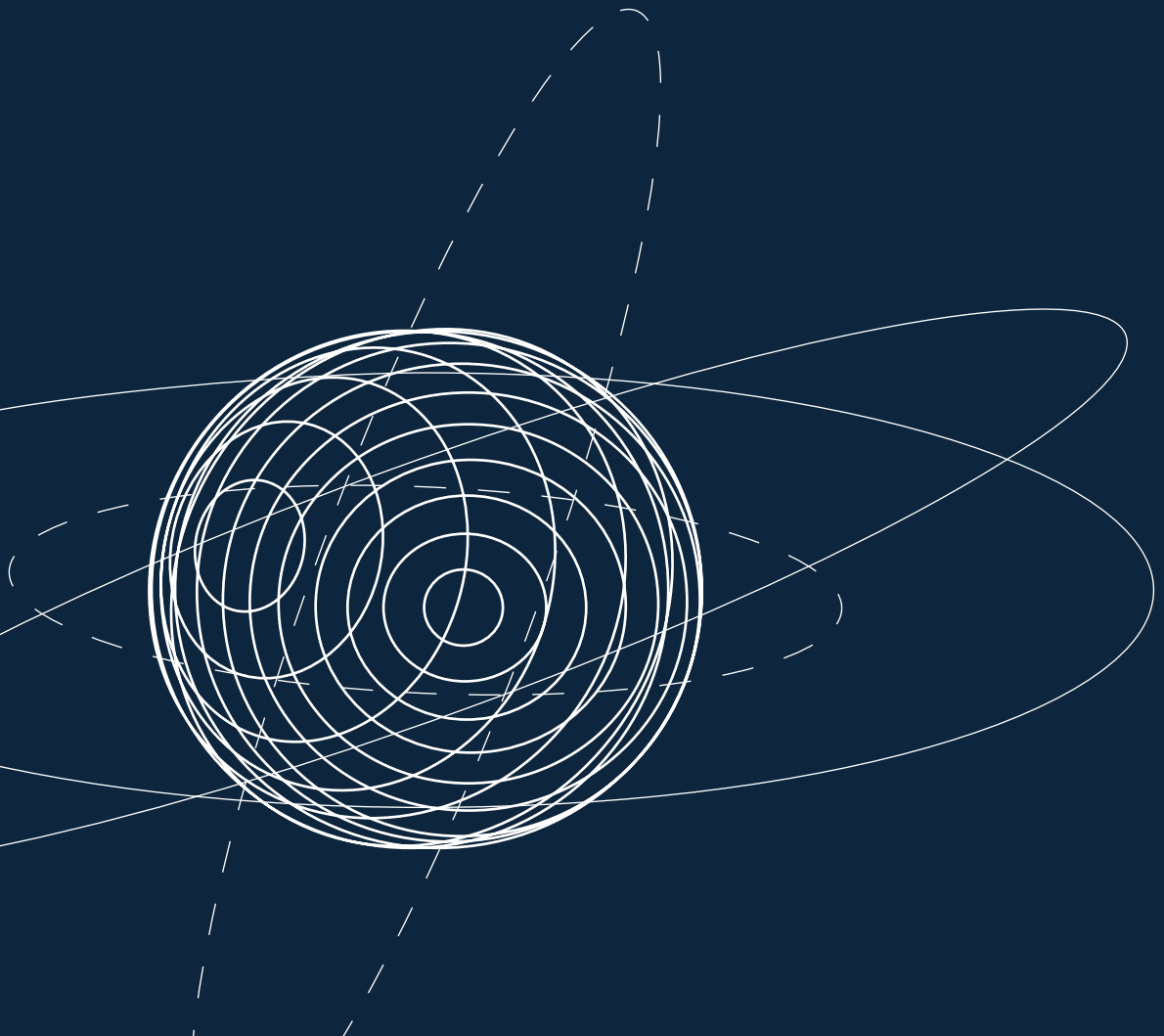
The study protocol is available at :
https://clinicaltrials.gov/ct2/show/NCT03599856



PART 5

A fully closed-loop automated CDS system

5.1



Current practice of closed-loop mechanical ventilation modes on intensive care units – a nationwide survey in the Netherlands

A.J.R. De Bie, E.F.E. Wenstedt, A.N. Roos, J.J.M. Verberne, H.H.M. Korsten, M.J. Schultz, A.J.G.H. Bindels

The Netherlands Journal of Medicine 2017

ABSTRACT

Background

The most recent modes for mechanical ventilation are closed-loop modes, which are able to automatically adjust certain respiratory settings. Although closed-loop modes have been investigated in various clinical trials, it is unclear to what extent these modes are actually used in clinical practice. The aim of this study was to determine closed-loop ventilation practice on intensive care units (ICUs) in the Netherlands, and to explore reasons for not applying closed-loop ventilation. Our hypothesis was that closed-loop ventilation is increasingly used.

Methods

A short survey was conducted among all non-paediatric ICUs in the Netherlands. Use of closed-loop modes was classified as frequently, occasionally or never, if respondents stated they had used these modes in the last week, in the last month/year, or never, respectively.

Results

The response rate of the survey was 82% (72 of 88). Respondents had access to a closed-loop ventilation mode in 58% of the ICUs (42 of 72). Of these ICUs, 43% (18 of 42) frequently applied a closed-loop ventilation mode, while 57% (24 of 42) never or occasionally used it. Reasons for not using these modes were lack of knowledge (40%), insufficient evidence reporting a beneficial effect (35%) and lack of confidence (25%).

Conclusion

This study does not support our hypothesis that closed-loop ventilation is increasingly used in the Dutch ICU setting. While industry continues to develop new closed-loop modes, implementation of these modes in clinical practice seems to encounter difficulties. Various barriers could play a role, and these all need attention in future investigations.

INTRODUCTION

Mechanical ventilation in intensive care unit (ICU) patients is a rapidly evolving field. Closed-loop ventilation modes are increasingly available, but it is uncertain to what extent they are used. Closed-loop ventilation modes automatically adjust certain respiratory settings based on digital algorithms and physiological inputs of the patient (e.g. pulse oximetry results, end-tidal CO_2 levels, and respiratory system resistance and compliance). Typical examples of closed-loop ventilation modes include Adaptive Support Ventilation (ASV®), INTELLiVENT®–ASV, SmartCare®/PS, Proportional Assist™ Ventilation (PAV™+), Neurally Adjusted Ventilatory Assistance (NAVA), Automode® and Mandatory Minute Ventilation (MMV).[1, 2] An international survey, published in 2011, reported that a majority of ICUs do not commonly use these modes, which was recently confirmed by a Ukrainian single-country study.[3, 4] Now, several years later, we hypothesise that closed-loop ventilation is increasingly applied. We performed a nationwide survey to determine closed-loop ventilation practice in ICUs in the Netherlands.

METHODS

A survey was conducted among all non-paediatric ICUs in the Netherlands. The study was registered at the Local Institutional Review Board of the Catharina Hospital, Eindhoven, the Netherlands. In September 2016, a representative ICU physician or nurse was identified for each ICU, and was then asked to participate in the survey. Participants could either answer the survey questions immediately by phone, or receive the survey by e-mail to be completed at a later stage. Two reminders were sent, one week and two weeks after the initial invitation. Non-responders were contacted again once more in November 2016.

The survey

The survey consisted of seven questions regarding the application of closed-loop modes (*Figure 1*). The use was classified as frequently, occasionally or never if respondents with a closed-loop ventilation mode had applied this mode at least once in the preceding week, month to year, or never, respectively. Reasons for not using closed-loop modes could be scored as 'lack of knowledge', 'insufficient evidence reporting a beneficial effect', or 'lack of confidence in the mode '. Respondents were also able to suggest additional reasons using an open field. An independent medical epidemiologist verified the methodological quality of the survey.

Analysis

The availability of a closed-loop mode and the frequency of use was analysed per ICU level. In the Netherlands, all Dutch ICUs have been classified from level 1, low level ICUs, to 3, high level ICUs, based on the ICU size, patient volume, ventilation days, and staffng.[5] Data were collected and entered into Microsoft Excel® version 14 (©2010 Microsoft Corporation). Categorical responses of questions were described as the proportion (percentage) of respondents selecting each response.

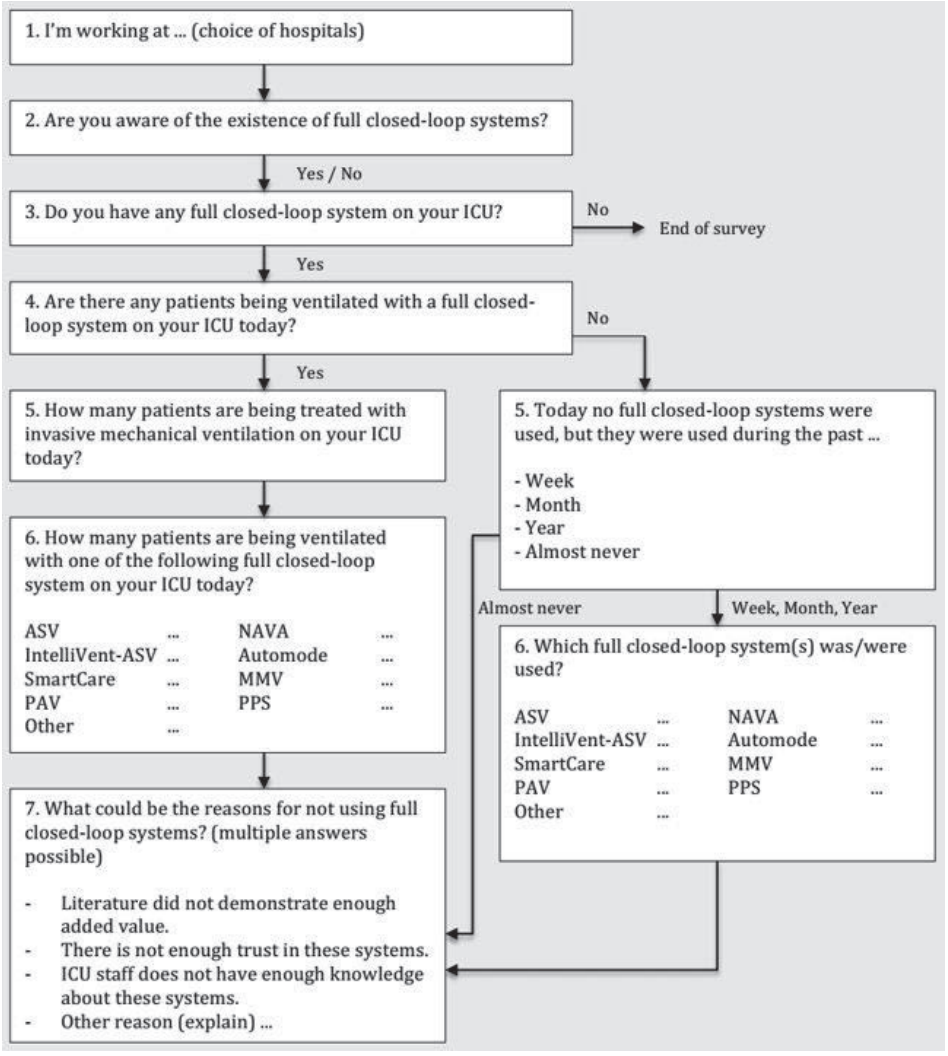


Figure 1. Flowchart-like survey for representatives of each non-pediatric ICU

RESULTS

The response rate of the survey was 82% (72 of 88). Respondents had access to a closed-loop ventilation mode in 58% of the ICUs (42 of 72) (*Figure 2*). Of these ICUs, 43% (18 of 42) frequently used a closed-loop ventilation mode, while 57% (24 of 42) occasionally or never used it (*Figure 3*). The majority of the frequent users were level 3 ICUs (50% vs. 11% and 39% level 1 and 2, respectively), whereas the majority of the occasional users consisted of level 1 ICUs (54% vs. 16% and 29% level 2 and 3, respectively; *Table 1 and 2*). The ICUs with INTELLiVENT®–ASV never classified the frequency of use as occasional or never. No other noticeable differences were observed between the frequent users and the occasional users with regard to the types of modes. On the day of the survey, 24% of the ICUs (10 of 42) reported having at least one patient on a closed-loop ventilation mode. These ICUs averagely ventilated 51% of their ventilated patients with a closed-loop mode.

Table 1. The availability of a closed-loop mechanical ventilation mode (yes/no) per ICU level (level 1, 2 or 3)

	Availability of a closed-loop mechanical ventilation mode:	
	Yes	**No**
Level 1	15 (35.71%)	15 (50.00%)
Level 2	11 (26.19%)	7 (23.33%)
Level 3	16 (38.10%)	8 (26.67%)
Total	42 (100%)	30 (100%)

Data on ICU levels was extracted from http://www.ziekenhuizentransparant.nl

Table 2. The frequency of use (frequently or occasionally/never) per ICU level (level 1, 2 or 3)

	Availability of a closed-loop mechanical ventilation mode:	
	Yes	**No**
Level 1	15 (35.71%)	15 (50.00%)
Level 2	11 (26.19%)	7 (23.33%)
Level 3	16 (38.10%)	8 (26.67%)
Total	42 (100%)	30 (100%)

Data on ICU levels was extracted from http://www.ziekenhuizentransparant.nl

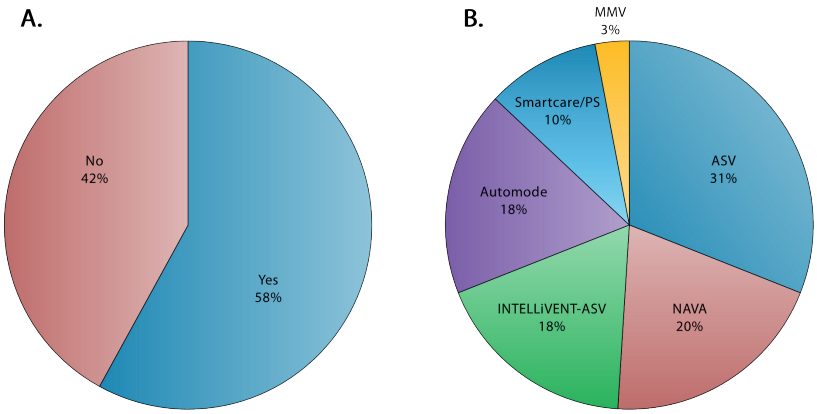


Figure 2. The percentage of ICUs that had access to a closed-loop ventilation mode (A), and the kinds of closed-loop ventilation modes that were used on these ICUs (B).
ASV®= adaptive support ventilation; NAVA = neurally adjusted ventilatory assistance and MMV = mandatory minute ventilation.

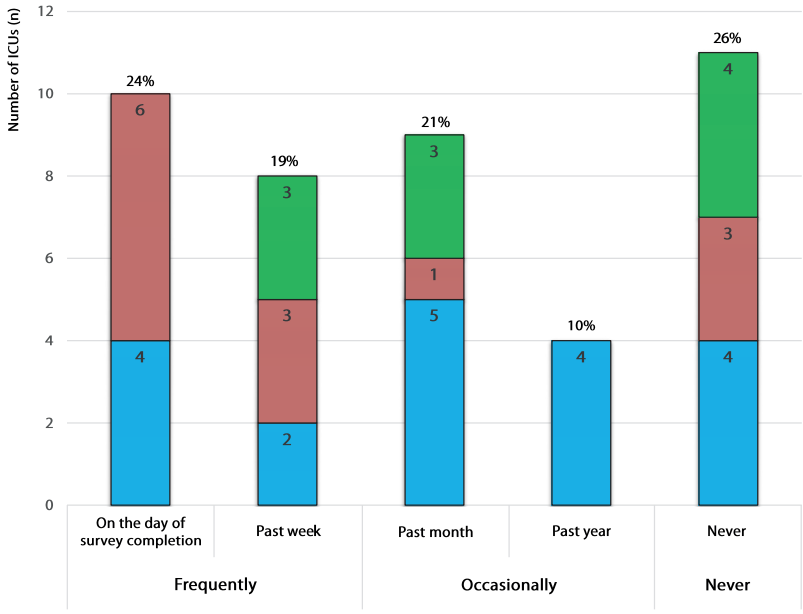


Figure 3. Frequency of use of closed-loop ventilation modes.
The last time a closed-loop ventilation mode was used on the different levels of ICUs with access to a closed-loop ventilation mode (n = 42). The use was classified as frequently, occasionally or never if respondents with a closed-loop ventilation mode had applied this mode at least once in the preceding week, month to year, or never, respectively.

Respectively 17, 14 and 11 ICUs with access to a closed-loop ventilation mode stated that reasons for not using this mode were lack of knowledge (41%), insufficient evidence reporting a beneficial effect (33%) and lack of confidence in the mode (26%) (*Figure* 4). Another 10% of these respondents mentioned that a perceived lack of control with the use of these modes might also play a role. With regard to (INTELLiVENT®–)ASV, 17% of the respondents expressed the concern that this mode selects higher tidal volumes than desired. Concerning NAVA, 7% of the respondents stated that the costs of the necessary disposables were a barrier for its use.

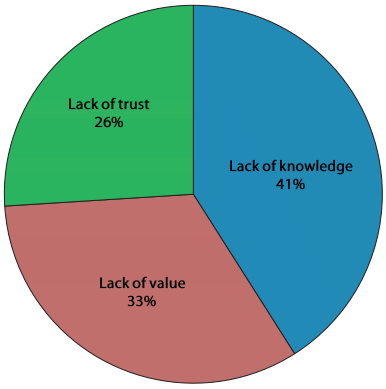


Figure 4. Reported reasons for not using closed-loop ventilation modes.
The percentage of participants, of ICUs with access to a closed-loop ventilation mode, that agreed with possible reasons for not using closed-loop ventilation modes.

DISCUSSION

The results of this survey echo those from the international European survey in 2011 and the Ukrainian survey in 2013,[3, 4] but do not support our hypothesis that closed-loop ventilation is increasingly used in the Dutch ICU setting. The most reported reason for resistance in our survey was 'lack of knowledge', which might be explained, at least in part, by a lack of experience and insufficient education, which are needed for acquiring knowledge and for successful implementation.[6] Both explanations depend on local manpower and on the case mix dependent culture of the ICU. Interestingly, this study shows that frequent users mainly consisted of high level ICUs, while occasional users were mostly lower level ICUs. One explanation could be that lower level ICUs have less staff and less time and means available for the introduction of new modes of ventilation, all leading to a more conservative culture.

The second-most mentioned reason for not using closed-loop ventilation modes was 'insufficient evidence reporting a beneficial effect'. While various studies have been performed, among which three recent meta-analyses, results are still not conclusive.[7-9] Additionally, in research, closed-loop mechanical ventilation modes are often grouped together, while these modes operate according to different techniques in order to achieve different goals for various indications. These considerations make the translation of research outcomes into clinical practice challenging, since it is uncertain to what extent this evidence can help the clinicians to choose for a specific closed-loop ventilation mode which best suits their specific case mix and local culture of the ICU.

The third reason for not using closed-loop modes was 'lack of confidence in the mode'. In highly controlled environments such as the ICU, where the staff attempt to control each parameter as much as possible, it can be difficult to entrust this process to a machine, also known as the 'black box effect'.[10] This could explain why 'lack of control' was added as an additional reason for not using closed-loop modes.

Our study has certain limitations. First, although we reached a high response rate, the design of the study potentially introduces selection bias as clinicians who use closed-loop ventilation modes may be more inclined to respond. This means that the implementation rate might be even lower in reality. Secondly, this survey did not register the version of the modes used, and some comments may be related to older versions of the ventilation modes. For instance, ASV, INTELLiVENT®-ASV and NAVA have had several updates, which improved safety (e.g., lower tidal volumes in the first two modes) and ease of use (e.g., less alarms in the last mode). Finally, this study does not provide a complete overview of all possible reasons that can influence the implementation of closed-loop modes. Many other possible contributing factors were not asked about in the survey, such as economic factors and long-term contracts with specific manufacturers.

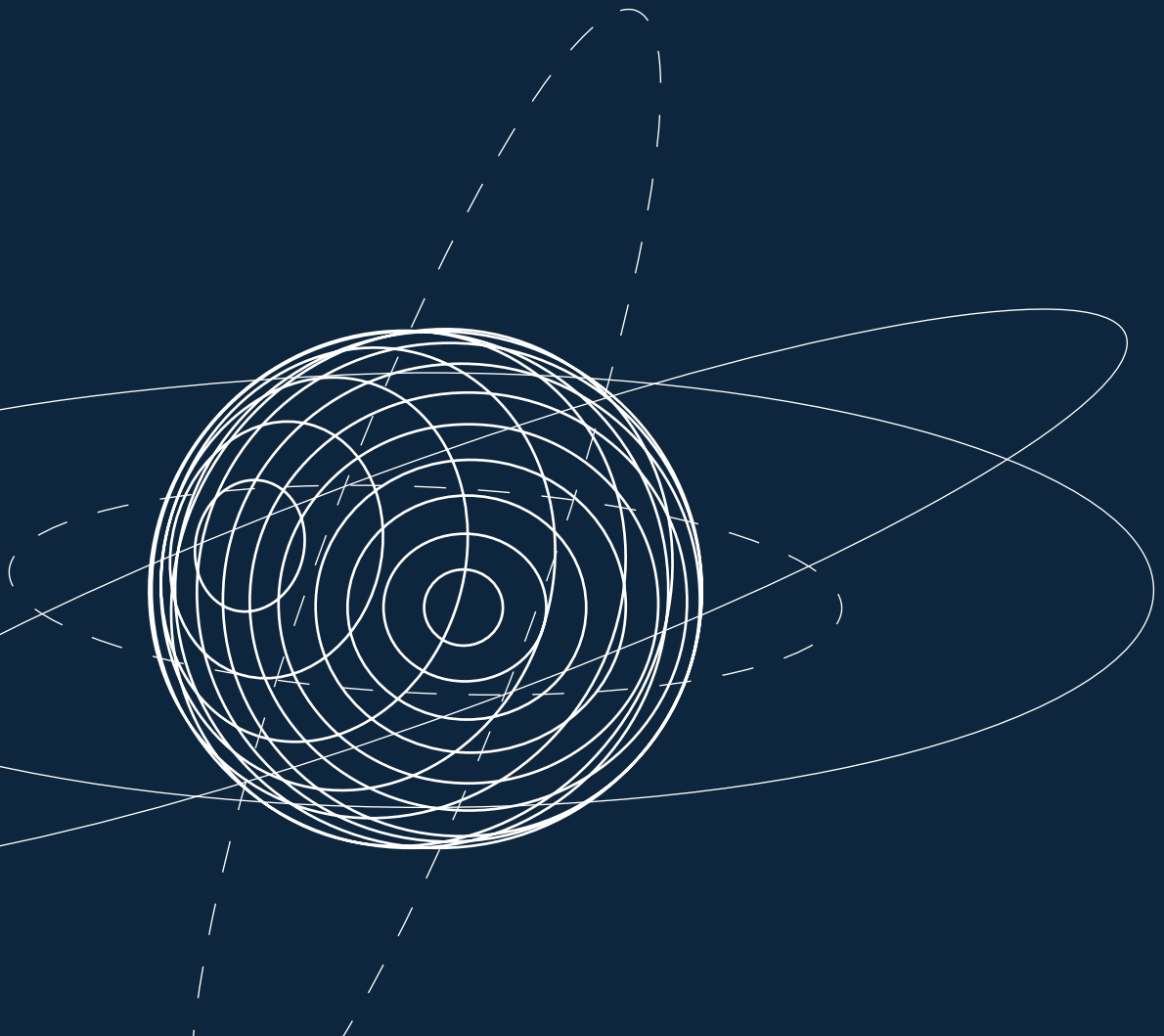
CONCLUSION

In conclusion, while industry continues to develop new closed-loop modes, implementation of these modes in clinical practice seems to encounter difficulties. Various barriers could play a role, and these all need attention in future investigations.

REFERENCES

1. Chatburn RL, Mireles-Cabodevila E. Closed-loop control of mechanical ventilation: Description and classifcation of targeting schemes. *Respir Care* 2011; **56**: 85-102.
2. Wysocki M, Jouvet P, Jaber S. Closed loop mechanical ventilation. *J Clin Monit Comput*. 2014; **28**:49-56.
3. Rose L, Blackwood B, Egerod I, et al. Decisional responsibility for mechanical ventilation and weaning: an international survey. Crit Care 2011; **15**:R295.
4. Pylypenko MM. Current practice of mechanical ventilation: Preliminary results of national survey in ukraine. *Intensive Care Med* 2013, **39**:S408.
5. Kluge GH, Brinkman S, van Berkel G, et al. The association between ICU level of care and mortality in the Netherlands. *Intensive Care Med* 2015; **41**:304-11.
6. Rubenfeld GD. Implementing effective ventilator practice at the bedside. *Curr Opin Crit Care* 2004; **10**:33-9.
7. Rose L, Schultz MJ, Cardwell CR, Jouvet P, McAuley DF, Blackwood B. Automated versus non-automated weaning for reducing the duration of mechanical ventilation for critically ill adults and children. *Cochrane Database Syst Rev* 2014; **6**: CD009235.
8. Burns KE, Lellouche F, Nisenbaum R, Lessard MR, Friedrich JO. Automated weaning and SBT systems versus non-automated weaning strategies for weaning time in invasively ventilated critically ill adults. *Cochrane Database Syst Rev* 2014; **9**: CD008638.
9. Burns KE, Lellouche F, Lessard MR, Friedrich JO. Automated weaning and spontaneous breathing trial systems versus non-automated weaning strategies for discontinuation time in invasively ventilated postoperative adults. *Cochrane Database Syst Rev* 2014; **2**: CD008639.
10. Tschopp M, Lovis C. Towards clinically available closed-loop systems for mechanical ventilation in the intensive care unit. *Yearb Med Inform* 2009: 75-80.

5.2



Fully automated closed-loop ventilation is safe and effective in post-cardiac surgery patients

A.J.R. De Bie, A.N. Roos, A.J.G.H. Bindels

Intensive Care Medicine 2014

INTRODUCTION

Dear Editor, A recent Cochrane review shows that automated ventilation, like assisted support ventilation (ASV), may reduce duration of weaning, ventilation, and ICU stay.[1] An extension of ASV is the fully automated closed-loop ventilating mode INTELLiVENT-ASV. Minute ventilation is not only automatically calculated on the basis of ASV's least work of breathing concept according to Otis,[2] but in combination with the patients end-tidal CO_2 ($EtCO_2$). And unlike ASV, it automatically adjusts FIO_2 and positive end-expiratory pressure (PEEP) on the basis of the ARDS Network PEEP-FIO_2 tables to maintain a target pulse oximetry.[3]

METHODS

We conducted a prospective noninferiority pilot study to determine the safety and efficacy of INTELLiVENT-ASV compared to ASV and our conventional ventilation (pressure controlled ventilation followed by pressure support ventilation) in patients weaning on a post-anaesthesia care unit (PACU). Included were low-risk post-cardiac surgery adults, suitable to wean on the PACU. Excluded were patients with a positive history of COPD Gold 3 or 4, lung surgery, and patients in shock. The ventilation mode could be changed when current ventilation was inefficient. The medical ethical committee approved the study and patients were excluded if they objected to use of their information.

RESULTS

In total 128 patients were included and divided into three groups, conventional ventilation (n = 49), INTELLiVENT-ASV (n = 53), and ASV (n = 26), based on the moment of admission at the PACU. Analysis of variance (ANOVA) between groups showed no statistically significant difference of age, BMI (kg m^{-2}), smokers, Euroscore, extracorporeal circulation time, and type of cardiac surgery. Ventilation-related safety issues requiring interventions were not observed in all groups. The number of interactions was statistically significantly lower in the INTELLiVENT-ASV group compared to the other groups (*Figure 1*). Mechanical ventilation time, the number of reintubations, and the amount of desaturations, defined as a SpO_2 lower than 85%, showed no statistically significant differences ($p>0.05$).

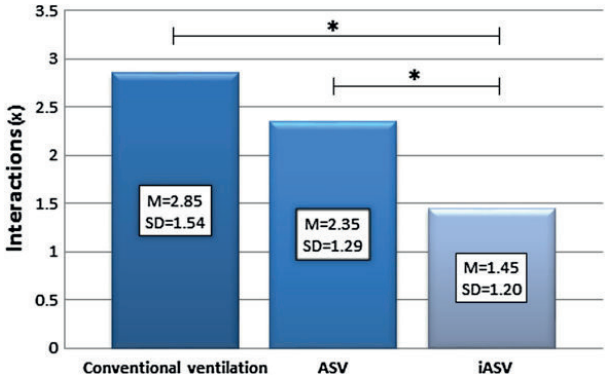


Figure 1. Representation of the number of interactions with the ventilator in the conventional ventilation group, ASV group, and INTELLiVENT-ASV group.
**=p<0.001. M=mean, SD=standard deviation.*

DISCUSSION

Fully automated closed-loop ventilation is able to mimic the dynamic process of human breathing by constantly adjusting ventilation and oxygenation depending on the individual demand. In our prospective trial we showed that full closed-loop ventilation with INTELLiVENT-ASV is a safe and effective mode to ventilate, oxygenate, and wean low-risk post-cardiac surgery patients.

The reduced number of interactions with the ventilator decreases workload, the risk of human errors, and may reduce inadequate ventilation time. This reduction could even be underestimated, because most physicians and nurses lacked confidence to extubate the patient directly from the new ventilation mode (INTELLiVENT-ASV).

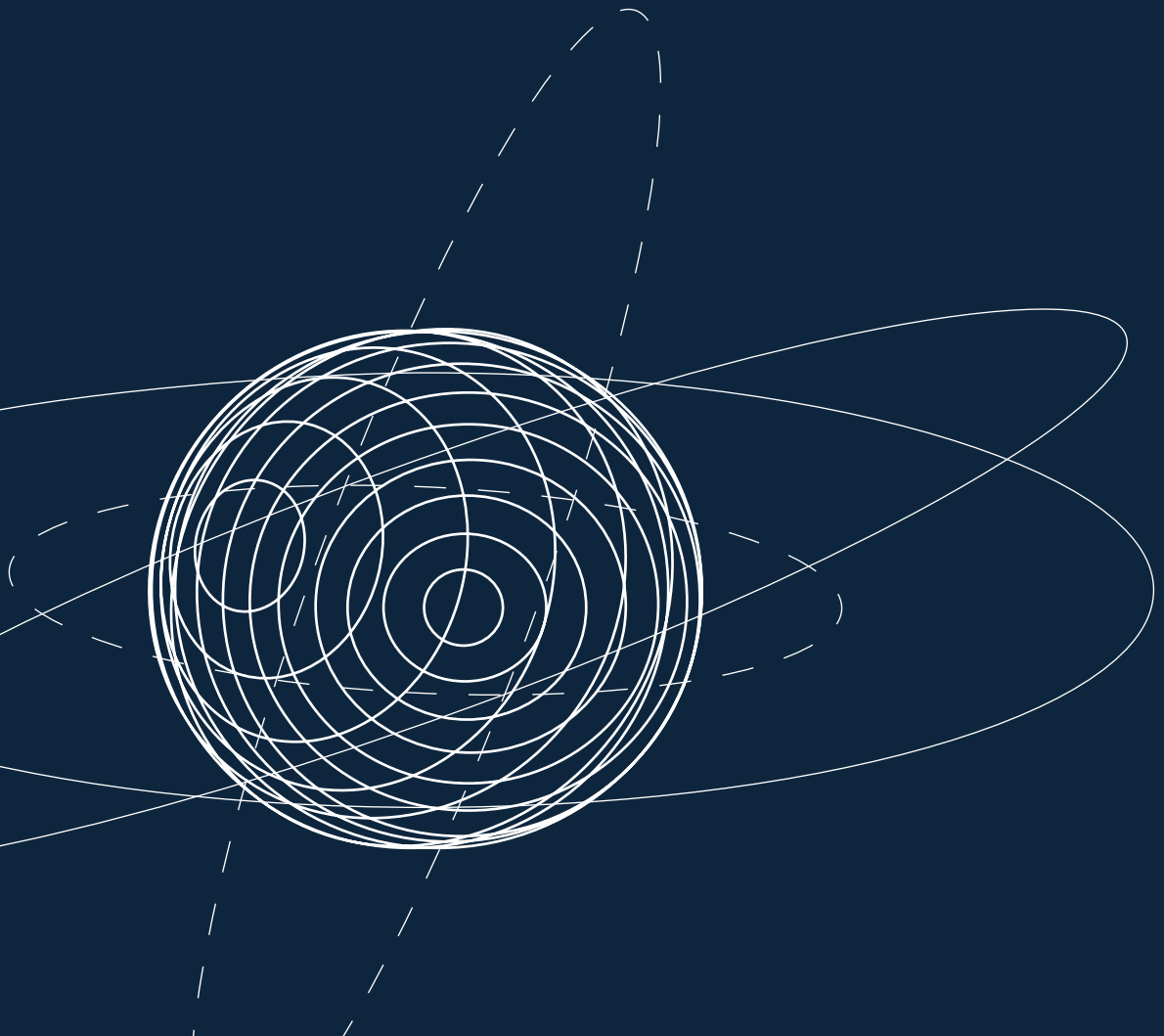
Our results were consistent with previous studies comparing INTELLiVENT-ASV with conventional ventilation modes.[4, 5] These studies even report a statistically significant higher percentage of acceptable and optimal ventilation time (99.5 % instead of 93%, $p<0.001$),[4] with statistically significant lower ventilating pressures, volumes, and FIO_2 in both low- and high-risk critically ill patients.[4, 5]

Our non-inferiority trial confirms that INTELLiVENT-ASV is as safe and efficient as conventional ventilation and ASV to ventilate and oxygenate weaning patients after cardiac surgery. However, more studies are needed in critically ill and postoperative patients to fully understand the clinical impact of fully closed-loop ventilation like INTELLiVENT-ASV.

REFERENCES

1. Rose L, Schultz MJ, Cardwell CR, Jouvet P, McAuley DF, Blackwood B. Automated versus non-automated weaning for reducing the duration of mechanical ventilation for critically ill adults and children. *The Cochrane database of systematic reviews* 2013: CD009235
2. Otis AB, Fenn WO, Rahn H. Mechanics of breathing in man. *Journal of applied physiology* 1950; **2**: 592-607
3. Brower RG, Lanken PN, MacIntyre N, et al. Higher versus lower positive end-expiratory pressures in patients with the acute respiratory distress syndrome. *The New England journal of medicine* 2004; **351**: 327-36
4. Lellouche F, Bouchard PA, Simard S, L'Her E, Wysocki M. Evaluation of fully automated ventilation: a randomized controlled study in post-cardiac surgery patients. *Intensive care medicine* 2013; **39**: 463-71
5. Arnal J-M, Garnero A, Novonti D, et al. Feasibility study on full closed-loop control ventilation (IntelliVent-ASV™) in ICU patients with acute respiratory failure: a prospective observational comparative study. *Critical care (London, England)* 2013; **17**: R196-R

5.3



Fully automated postoperative ventilation in cardiac surgery patients: a randomised clinical trial

A.J.R. De Bie, A. Serpa Neto, D.M. van Meenen, R.A. Bouwman, A.N. Roos, J.R. Lameijer, H.H.M. Korsten, M.J. Schultz, A.J.G.H. Bindels

British journal of anaesthesia 2020

ABSTRACT

Background

Ensuring that lung-protective ventilation is achieved at scale is challenging in perioperative practice. Fully automated ventilation may be more effective in delivering lung-protective ventilation. Here, we compared automated lung-protective ventilation with conventional ventilation after elective cardiac surgery in haemodynamically stable patients.

Methods

In this single-centre investigator-led study, patients were randomly assigned at the end of cardiac surgery to receive either automated (adaptive support ventilation) or conventional ventilation. The primary endpoint was the proportion of postoperative ventilation time characterised by exposure to predefined optimal, acceptable, and critical (injurious) ventilatory parameters in the first three postoperative hours. Secondary outcomes included severe hypoxaemia (SpO_2 <85%) and resumption of spontaneous breathing. Data are presented as mean (95% confidence intervals [CIs]).

Results

We randomised 220 patients (30.4% females; age: 62–76 yr). Patients randomised to automated ventilation (n=109) spent a 29.7% (95% CI: 22.1–37.4) higher mean proportion of postoperative ventilation time receiving optimal postoperative ventilation after surgery (P<0.001) compared with patients receiving conventional postoperative ventilation (n=111). Automated ventilation also reduced the proportion of postoperative ventilation time that patients were exposed to injurious ventilatory settings by 2.5% (95% CI: 1–4; P=0.003). Severe hypoxaemia was less likely in patients randomised to automated ventilation (risk ratio: 0.26 [0.22–0.31]; P<0.01). Patients resumed spontaneous breathing more rapidly when randomised to automated ventilation (hazard ratio: 1.38 [1.05–1.83]; P=0.03).

Conclusion

Fully automated ventilation in haemodynamically stable patients after cardiac surgery optimised lung-protective ventilation during postoperative ventilation, with fewer episodes of severe hypoxaemia and an accelerated resumption of spontaneous breathing.

INTRODUCTION

High tidal volume during postoperative ventilation is a risk factor for organ dysfunction, pulmonary complications, and prolonged ICU stay in cardiac surgery patients.[1] High PEEP may prevent postoperative complications and shortens ICU stay in cardiac surgery patients who present with hypoxaemia upon arrival in the ICU.[2] Physiological and clinical studies suggest that arterial hyperoxia and also hypoxia are better avoided in ventilated patients, as both have an association with mortality.[3–9]

Automated modes of ventilation are increasingly becoming available for clinical use.[10,11] The common goals of automated ventilatory modes are to tailor ventilator settings to patient's needs, facilitate earlier recognition of the ability to breathe spontaneously with subsequent smooth weaning from the ventilator,[12–14] and deliver lung-protective ventilator settings.[11,15] INTELLiVENT-Adaptive Support Ventilation (ASV) is a fully automated, or closed-loop, ventilation mode that consists of pressure-controlled ventilation or pressure support ventilation depending on a patient's respiratory activity. In fully automated ventilation mode, tidal volume, pressure levels (including PEEP), minute ventilation, and the oxygen fraction in inspired air are controlled solely by the ventilator.[16,17]

Previous studies have shown INTELLiVENT-ASV to be capable of applying ventilation with safe ventilator settings in critically ill patients.[18–27] The aim of the current study was to compare INTELLiVENT-ASV with conventional ventilation during postoperative ventilation after uncomplicated cardiac surgery. We hypothesised that fully automated mode of ventilation would be more likely to deliver lung-protective ventilation during weaning after cardiac surgery.

METHODS

Study design and oversight

The Postoperative INTELLiVENT-ASV Ventilation study was an investigator-initiated, single-centre, parallel-group, randomised clinical trial, conducted at the ICU of a tertiary teaching hospital in Eindhoven, Netherlands. The manufacturer of the ventilator was not involved. The study protocol was approved by the local Institutional Review Board (R16.054) and registered at ClinicalTrials.gov (study identifier: NCT03180203). A statistical analysis plan was constructed before cleaning and closing the database; the

final plan and a table describing changes to the original analysis plan are available in the Supplementary material. Written informed consent was obtained from all individual participants before surgery.

Inclusion criteria

Patients were eligible if they were scheduled for elective cardiac surgery requiring postoperative invasive ventilation in the ICU.

Exclusion criteria

Before surgery, patients were excluded if aged <18 yr, BMI >35 kg m^{-2}, have a history of pneumonectomy or lobectomy, presence of chronic obstructive pulmonary disease (COPD) (Global Initiative for Chronic Obstructive Lung Disease Class III or IV), or if they were already enrolled in another interventional trial. After cardiac surgery, patients were ineligible if extracorporeal support was required after surgery, or if they were deemed by the attending clinician to be haemodynamically unstable. Fast-track cardiac surgery patients were also ineligible as they were planned to receive postoperative ventilation in the PACU where the automated mode was not available.

Randomisation and masking

Patients were randomised in a 1:1 ratio to fully automated ventilation (the 'automated group') or conventional ventilation ('conventional group') before the start of surgery. Local investigators performed randomisation with a web-based randomisation programme that used random block sizes. Physicians and nurses caring for the patients in the ICU could not be blinded because of the nature of the intervention. The investigators who performed the analyses (AJRDB and ASN) and the radiologist (JRL), though, remained blind for randomisation at all times.

Perioperative care

Standardised perioperative care was followed according to local guidelines. Typically, one board-certified ICU nurse cared for a maximum of two patients. Nurses were responsible for adjusting ventilator settings; doctors could be consulted at all times. Arterial blood gases were performed regularly. Extubation criteria were similar for both groups and followed the local guideline. For additional details on standard care, see Supplementary data.

Study interventions

The same type of ventilator (Hamilton-S1; Hamilton Medical, Rhäzüns, Switzerland) was used for all patients. All attending ICU nurses and physicians were trained and qualified to use this ventilator and the INTELLiVENT-ASV and the volume-controlled ventilation mode.

In the automated group, ventilation started with volume-controlled ventilation. INTELLiVENT-ASV was initiated as soon as the first blood gas analysis was available, typically within 15 minutes after arrival in the ICU. After initiating INTELLiVENT-ASV, minute ventilation VT, pressure levels (including PEEP), and FIO_2 were automatically adjusted by the ventilator to provide invasive ventilation within appropriate ranges of $EtcO_2$ and SpO_2. Thus, neither VT and minute volume nor PEEP and FIO_2 were to be adjusted by the attending ICU nurse or doctor. For additional details on settings with INTELLiVENT-ASV, see the Supplementary material.

In the conventional group, ventilation also started with volume-controlled ventilation, and pressure support was initiated as soon as the patient was able to trigger the ventilator, which was typically tested every 15 minutes after cessation of postoperative sedation. VT, maximum airway pressure (Pmax), and ventilatory frequency (VF) were manually titrated to have VT ≤7 ml kg^{-1} predicted body weight (PBW); Pmax <30 cm H_2O. Ventilatory frequency was titrated to have $EtcO_2$ between 4.7 and 6.4 kPa. PEEP and FIO_2 were titrated using a low PEEP–FIO_2 table to have SpO_2 stay between 93% and 98%.[16] For additional details on settings with conventional, see the Supplementary material.

Data collection

'Breath-by-breath' ventilation data were collected using a StudyRecorder (version 1.5; Hamilton Medical) connected to study ventilators. Every 30 minutes, an inspiratory hold was performed to measure plateau pressure and an expiratory hold to measure total PEEP. Driving pressure and mechanical power of ventilation were calculated using the following formulae:

- Driving pressure = plateau pressure – PEEP
- Mechanical power (J min^{-1})=0.098 * VT (L) * VF * (maximum airway pressure–driving pressure * 0.5)

Inspiratory and expiratory holds were not performed during spontaneous breathing, meaning that driving pressure could only be estimated and mechanical power only be

calculated when a patient was receiving pressure controlled with INTELLiVENT-ASV in the automated group or volume-controlled ventilation in the control group (additional details are provided in Supplementary material).

Primary outcome

The primary outcome was the proportion of time spent in three predefined and previously used zones of ventilation in the first 3 hours of postoperative ventilation (*Table 1*).[23]

Table 1. Zones of ventilation used to define the outcomes of the study*

	Optimal Zone	Acceptable Zone	Critical Zone
Tidal volume, ml kg-1 PBW	≤ 8	8 – 12	> 12
	AND	AND/OR	OR
Maximum airway pressure, cmH_2O	< 31	31 – 36	≥ 36
	AND	AND/OR	OR
$etCO_2$, kPa	4.0 – 6.1	3.3 – 4.0 OR 6.1 – 6.8	< 3.3 OR ≥ 6.8
	AND	AND/OR	OR
SpO_2, %	93 – 98 OR ≥ 93 if FIO_2 ≤ 40%	≥ 98 OR 85 – 93	< 85
Definition	If any present critical zone	If not in the optimal zone and none of the critical zone is present acceptable zone	All must be present optimal zone
Missing	If all parameters are missing, zone is missing If parameters are missing but one is available and it is in the critical zone, zone is defined as critical If parameters are missing but one is available and it is NOT in the critical zone, zone is defined as missing		

*PBW: predicted body weight; $etCO_2$: end-tidal carbon dioxide; SpO_2: oxygen pulse oximetry; FIO_2: inspired fraction of oxygen; * adapted from Lellouche F, Bouchard PA, Simard S, L'Her E, Wysocki M. Evaluation of fully automated ventilation: a randomized controlled study in post-cardiac surgery patients. Intensive Care Med 2013;39:463-71.*

Secondary outcomes

We assessed the following secondary endpoints:

- Proportion of breaths within each predefined ventilatory zone in the first 3 hours of postoperative ventilation; outcomes regarding proportion of time spent in the three zones of ventilation were also reanalysed using the complete ventilation time instead of the first three postoperative hours

- Incidence of severe hypoxaemia (percentage of breaths with SpO_2 <85% if SpO_2 quality index was ≥50%)
- Time to spontaneous breathing (time from ICU admission until more than or equal to five consecutive spontaneous breaths)
- Duration of postoperative ventilation
- Duration of weaning (time from cessation of sedatives until tracheal extubation)
- Proportion of failed extubations (re-intubation within 48 hours after extubation, excluding patients re-intubated for re-sternotomy) and development of postoperative pulmonary complications (composite of pneumonia, pneumothorax, or severe atelectasis)
- ICU length of stay and readmission
- ICU and 30-day mortality

Statistical analysis

All analyses were performed in a modified intention-to-treat population. Reasons for exclusion until ICU admission (i.e. after randomisation) were haemodynamic instability at the end of surgery, with or without need to continue extracorporeal support after surgery, and incidentally the unavailability of a ventilator that could provide the fully automated ventilation mode. In the per-protocol analysis, patients who had one or more major protocol violations were excluded. Details are provided in the Supplementary material.

Descriptive data are reported as numbers and percentages, means (standard deviation), or medians (inter-quartile ranges). Comparison of ventilatory parameters between groups over time was done using mixed-effect longitudinal models with random intercepts for patients. For analysis of the primary outcome, Student's t-test was used with 95% confidence interval (CI), and results are presented as mean differences (MDs). For outcomes assessing proportions of breaths, the denominator was the total number of breaths. Secondary binary outcomes were assessed with risk ratio and 95% CI calculated with Wald likelihood ratio approximation test and χ^2 tests for hypothesis testing.

The effects of the intervention on time to spontaneous breathing, duration of weaning and ventilation, and 30-day mortality were assessed using Kaplan–Meier survival curves and reported as hazard ratios with 95% CI calculated from a Cox proportional hazard model. The Schoenfeld residuals against the transformed time were used to test the proportional hazard assumptions. Survival time was calculated from time of randomisation until time of the outcome. The effect of the intervention on ICU length of stay was estimated with generalised linear models using inverse Gaussian distribution.

In pre-specified exploratory analyses, the effects of automated ventilation on the proportions of time spent in critical zone were investigated in subgroups based on the following patient categories: (1) according to intraoperative ventilation time (shorter or longer than the median) and (2) according to PaO_2/FIO_2 (below or above the median at ICU admission). The effects in the subgroups were evaluated by generalised linear models considering Gaussian distribution. Although reported in the statistical analysis plan, an exploratory analysis according to duration of postoperative ventilation was not performed as this characteristic might be influenced by the intervention.

In one post hoc analysis, the ventilation zones were based on the four individual elements (i.e. maximum airway pressure, tidal volume, $EtCO_2$, and SpO_2). In a second post hoc analysis, groups were compared with respect to proportion of breaths: (1) with hyperoxia (SpO_2 >97%), hypoxaemia (SpO_2 <90%), and normoxia (SpO_2 between 90% and 97%); (2) with a Pmax of ≤30 cm H_2O; and (3) with a driving pressure ≤15 cm H_2O. Details are provided in the Supplementary material.

All analyses were performed using R software, version 3.4.1 (R Core Team, Vienna, Austria). Significance level for all outcomes was 0.05, without adjustment for multiple comparisons. All secondary outcomes and analyses were exploratory. Reported P-values are two sided, and because the amount of missing data is negligible, only complete case analysis was carried out.

Sample size calculation

The study sample size was calculated using G*power (version 3.1.9.2; Kiel, Germany). We estimated that a sample size of 196 patients would provide 95% power to detect a difference of 3% of ventilation time in the critical ventilation zone, based on findings in a previous study and an estimated baseline standard deviation of 2.5% of ventilation time, with a Type I error of 5% and corrected for dropouts.[23]

RESULTS

Patient characteristics

From May 20, 2017 to April 19, 2018, 712 patients were screened (*Figure 1*). Of 220 randomised patients, 109 were allocated to the automated group and 111 to the conventional group. Baseline characteristics and dosages of peri- and postoperative i.v. sedative and analgesic medications were similar between the study groups (*Table 2*; Supplementary Table S1). Fully automated ventilation started 9 (4–21) min after arrival at the ICU, which was attributable to time needed to obtain the results of the first

blood gas analysis required for programming of fully automated ventilation. Ventilator characteristics and initial arterial blood gas analyses are shown in Supplementary Tables S2 and S3 and Supplementary Figures S2–S4.

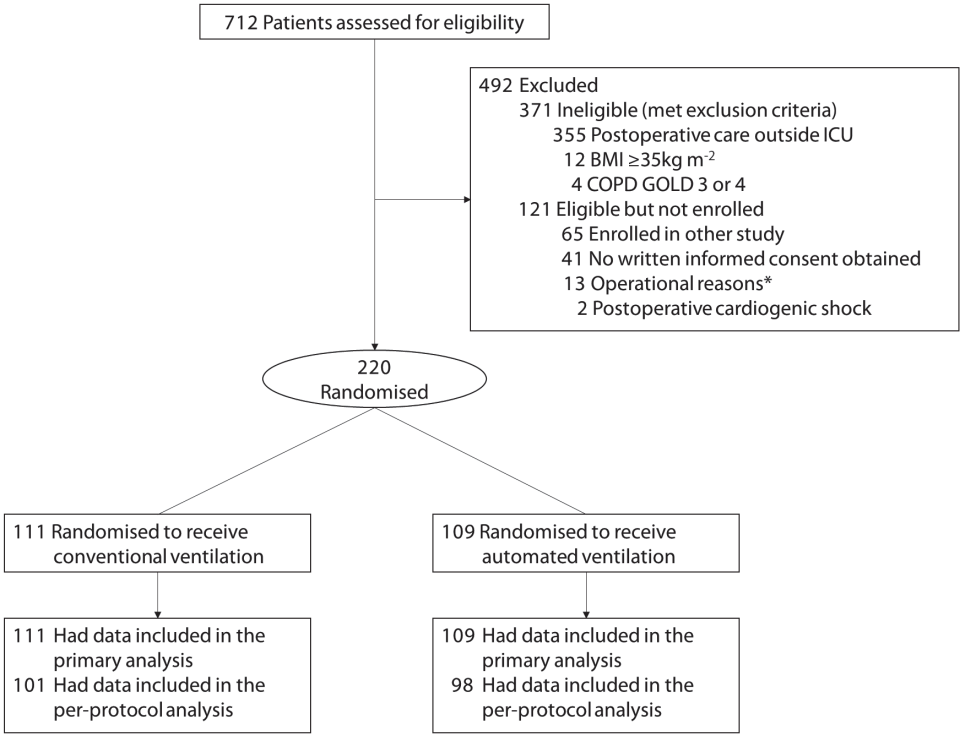


Figure 1 Flow of patients in the trial

Abbreviations: COPD, chronic obstructive pulmonary disease; GOLD, Global Initiative for Chronic Obstructive Lung Disease.

**No study ventilator available*

Table 2. Baseline characteristics of the included patients

	Fully automated ventilation (*n* = 109)	**Conventional ventilation (*n* = 111)**
Age, years	70 (62 – 76)	70 (63 – 76)
Male gender	73 (67.0)	80 (72.1)
PBW, kg	66.0 (59.7 – 75.1)	68.3 (61.0 – 73.3)
BMI, kg m^{-2}	26.0 (24.2 – 29.2)	26.5 (24.5 – 29.0)
SAPS II	31 (29 – 39)	33 (28 – 39)
APACHE IV	41 (33 – 49)	38 (32 – 48)
EuroSCORE II	1.6 (1.0 – 3.6)	1.6 (1.0 – 2.8)
Edmonton Frail Scale	3 (2 – 4)	3 (2 – 5)
Smoking		
No	39 (35.8)	47 (42.3)
Current	19 (17.4)	20 (18.0)
Former	51 (46.8)	44 (39.6)
Use of alcohol	65 (59.7)	71 (64.0)
COPD	7 (6.4)	12 (10.8)
Asthma	8 (7.3)	6 (5.4)
OSA	7 (6.4)	10 (9.0)
Diabetes mellitus	28 (25.7)	15 (13.5)
Hypertension	68 (62.4)	65 (58.6)
CVD or TIA	14 (12.9)	17 (15.3)
Heart failure	95 (87.9)	94 (85.4)
NYHA classification		
I	19 (17.6)	16 (14.5)
II	51 (47.2)	60 (54.5)
III	24 (22.2)	17 (15.5)
IV	1 (0.9)	1 (0.9)
Peripheral artery disease	19 (17.4)	14 (12.6)
Chronic kidney disease, %	23 (21.1)	28 (25.2)
LVEF	50 (35 – 56)	54 (45 – 60)
Right ventricular function		
Good	103 (97.2)	95 (94.1)
Moderate	2 (1.9)	5 (5.0)
Poor	1 (0.9)	1 (1.0)
Aortic valve disease		
None	54 (49.5)	47 (42.3)
Moderate insufficiency	5 (4.6)	5 (4.5)
Severe insufficiency	10 (9.2)	7 (6.3)
Moderate stenosis	6 (5.5)	4 (3.6)
Severe stenosis	34 (31.2)	48 (43.2)

Table 2. Continued

	Fully automated ventilation (*n* = 109)	**Conventional ventilation (*n* = 111)**
Mitral valve disease		
None	71 (65.1)	69 (62.2)
Moderate insufficiency	8 (7.3)	10 (9.0)
Severe insufficiency	29 (26.6)	29 (26.1)
Severe stenosis	1 (0.9)	3 (2.7)
Tricuspid valve disease		
None	96 (88.1)	90 (81.1)
Moderate insufficiency	8 (7.3)	14 (12.6)
Severe insufficiency	5 (4.6)	7 (6.3)
Preoperative use of Levosimendan	4 (3.7)	5 (4.5)
Type of surgery		
CABG	8 (7.3)	16 (14.4)
Valve surgery	48 (44.0)	47 (42.3)
CABG + Valve surgery	30 (27.5)	36 (32.4)
Off-pump CABG	14 (12.8)	2 (1.8)
Aortic repair	8 (7.3)	10 (9.0)
Myxoma excision	1 (0.9)	0 (0.0)
Duration of extracorporeal circulation, minutes	114 (87 – 157)	106 (77 – 145)
Duration of aortic occlusion, minutes	77 (57 – 109)	71 (53 – 97)
Peri-operative use of sedatives and analgesia		
Etomidate (mg)	50 (50 – 70)	50 (50 – 50)
Rocuronium (mg)	200 (200 – 200)	200 (200 – 200)
Propofol (mg)	1437 (1135 – 1760)	1404 (1130 – 1782)
Midazolam (mg)	2.5 (0 – 5)	5 (0 – 5)
Opiates		
Morphine (mg)	25 (25 – 25)	25 (25 – 25)
Alfentanil (mg)	1233 (955 – 1533)	1189 (1015 – 1545)
Sufentanil (mcg)	0 (0 – 0)	0 (0 – 0)
First postoperative level of CK–MB, U L^{-1}	59.5 (40.0 – 96.5)	58.0 (42.0 – 86.0)

Data are median (25 – 75% quartile) or No (%)

PBW: predicted body weight; BMI: body mass index; SAPS: Simplified Acute Physiology Score; APACHE: Acute Physiology and Chronic Health Evaluation; COPD: chronic obstructive pulmonary disease; OSA: obstructive sleep apnea; CVD: cerebrovascular disease; TIA: transient ischemic attack; NYHA: New York Heart Association; LVEF: left-ventricular ejection fraction; CABG: coronary artery bypass graft; CK-MB: Creatine Kinase MB Isoenzyme

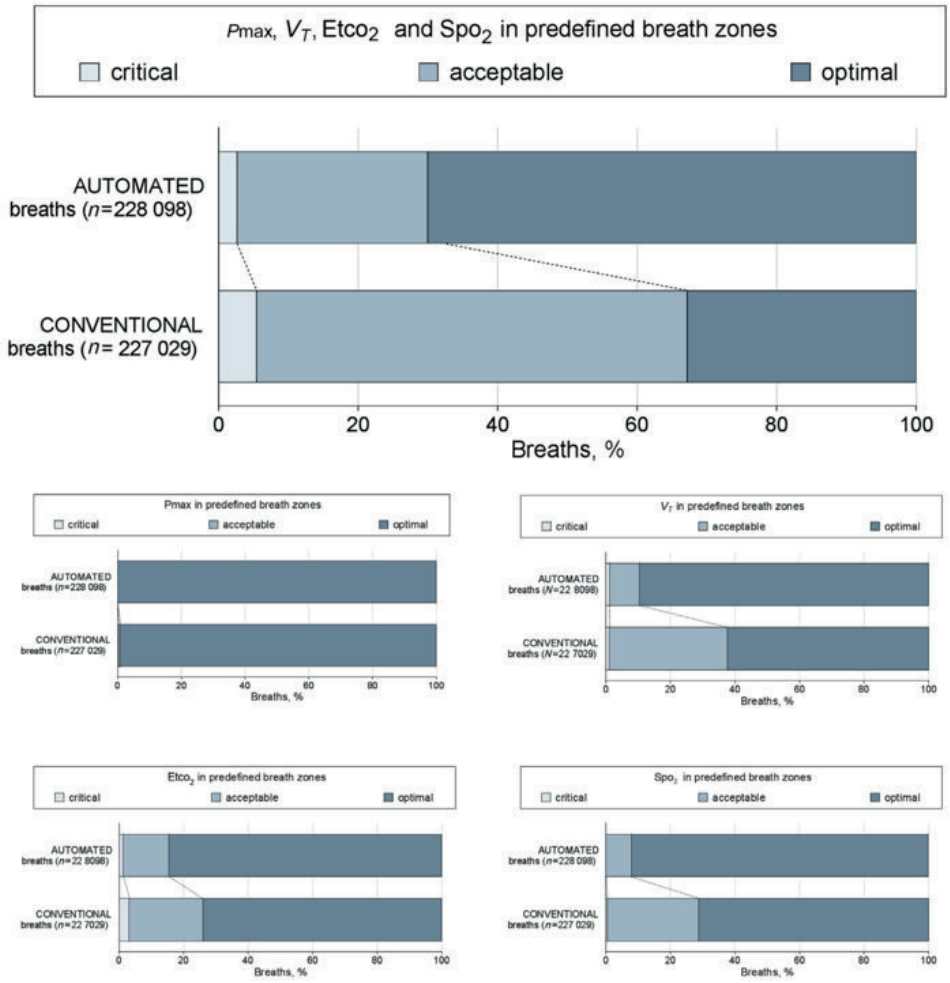


Figure 2. Percentage of breaths in predefined zones of ventilation.

Primary outcome

Patients in the fully automated group had a higher proportion of breaths in the optimal zone (*Figure* 2), as illustrated by heat maps of ventilation in consecutive blocks of 15 minutes for the first 3 hours of postoperative invasive (*Figure 3*; Supplementary Figures S5–S8). Patients in the automated group spent more time in optimal zones (55.2% [28.0]) compared with 25.5% [29.3] for conventionally ventilated patients (MD: 29.7; 95% CI: 22.1–37.4; P<0.001; *Table 3*). Patients in the automated group spent less time in the critical ventilation zone (0.5 [2.9%]) compared with 3.0 (8.3%) for conventionally ventilated patients (MD: 2.5% [95% CI: 0.8–4.1]; P=0.003) (*Table 3*). Accordingly, less time

was spent in the acceptable zone (automated ventilation: 16.7% [16.7]) compared with 50.0% [34.0] for conventionally ventilated patients (MD: –33.2 [95% CI: –40.4 to –26.1]; P<0.001). Reanalysis taking into account the absolute period of ventilation required for each patients before liberation from the ventilator gave similar results (Supplementary Tables S4 and S6; Supplementary Figure S9).

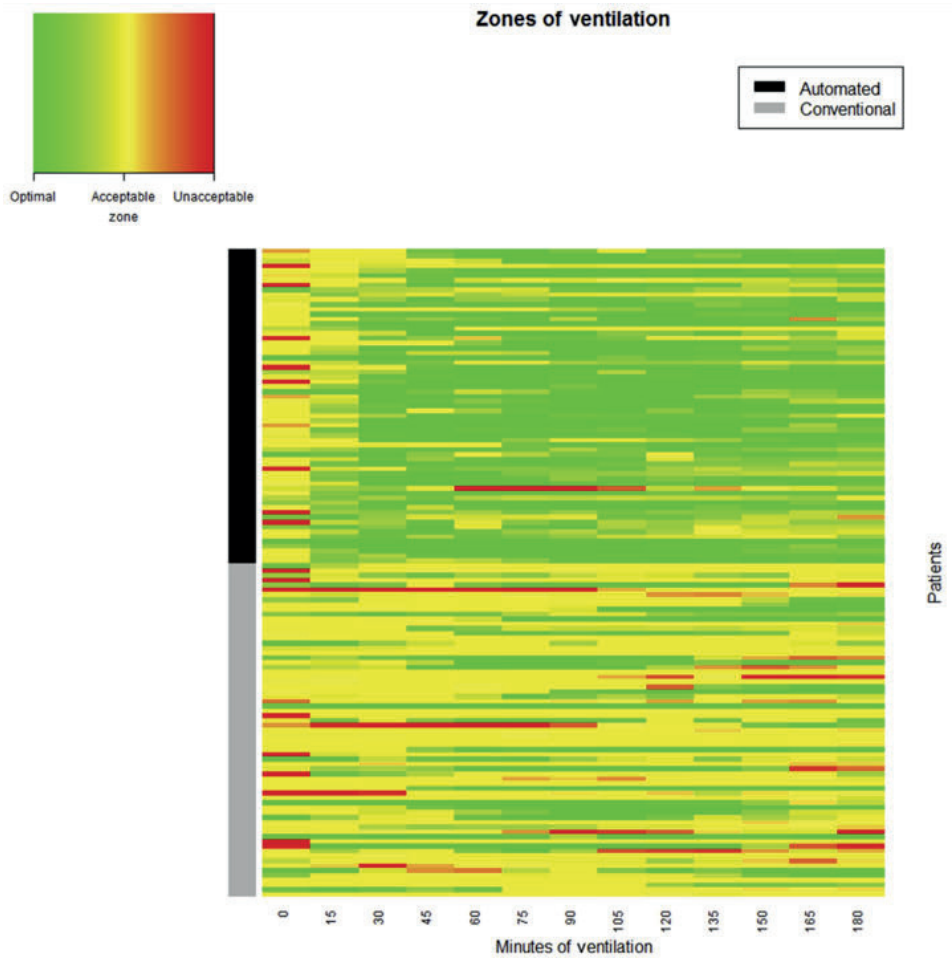


Figure 3. Heat map showing the ventilation zones every 15 minutes after randomisation.
For the heat map construction, each breath was assigned to the corresponding ventilation zone and was given a numeric value as follows: (1) for a breath in the optimal zone, (2) for a breath in the acceptable zone, and (3) for a breath in the unacceptable zone. Then, all value of the breaths recorded within every 15 minutes were summarised using the mean of these values. The default colour gradient sets the lowest mean value to green (Optimal), the highest value to a bright red (Unacceptable), and mid-range values to yellow (Acceptable), with a corresponding transition between these extremes.

Table 3. Co-primary and secondary outcomes

	Fully automated ventilation (n = 109)	Conventional ventilation (n = 111)	Effect Estimate (95%–CI)	p value
Co-primary outcomes				
Percentage of time in the optimal zone*	55.2 ± 28.0	25.5 ± 29.3		< 0.001[f]
Median (IQR)	61.2 (32.4 – 78.9)	15.3 (0.0 – 49.0)	29.7 (22.1 to 37.4)[a,f]	
Percentage of time in the acceptable zone*	16.7 ± 16.7	50.0 ± 34.0		< 0.001[f]
Median (IQR)	13.0 (4.6 – 22.9)	46.1 (19.0 – 83.6)	-33.2 (-40.4 to -26.1)[a,f]	
Percentage of time in the critical zone*	0.5 ± 2.9	3.0 ± 8.3		0.003[f]
Median (IQR)	0.0 (0.0 – 0.3)	0.3 (0.0 – 1.2)	-2.5 (-4.1 to -0.8)[a,f]	
Secondary outcomes				
Percentage of breaths in the optimal zone**	159,643 / 228,098 (70.0)	74,537 / 227,021 (32.8)	2.20 (2.18 to 2.21)[b]	<0.01
Percentage of breaths in the acceptable zone**	62,359 / 228,098 (27.3)	140,000 / 227,021 (61.7)	0.46 (0.46 to 0.47)[b]	<0.01
Percentage of breaths in the critical zone**	6,096 / 228,098 (2.7)	12,484 / 227,021 (5.5)	0.64 (0.63 to 0.65)[b]	<0.01
Time until spontaneous breathing, minutes££	164.0 ± 103.4	211.7 ± 169.5		
Median (IQR)	142 (90 – 219)	162 (114 – 260)	1.38 (1.05 to 1.83)[c]	0.02
Duration of weaning, minutes£	222.7 ± 328.2	265.7 ± 425.4		
Median (IQR)	136 (73 – 241)	157 (75 – 323)	1.16 (0.88 to 1.53)[c]	0.29
Duration of ventilation, minutes****	448.3 ± 1085.2	430.5 ± 457.2		
Median (IQR)	279 (195 – 412)	304 (204 – 477)	1.17 (0.89 to 1.54)[c]	0.26
Proportion of failed extubations***	2 / 103 (1.9)	2 / 105 (1.9)	1.01 (0.40 to 2.71)[b]	0.99
Incidence of postoperative pulmonary complications	87 / 109 (79.8)	96 / 111 (86.5)	0.80 (0.59 to 1.09)[b]	0.18
Pneumonia	7 / 109 (6.4)	7 / 111 (6.3)	1.01 (0.59 to 1.73)[b]	0.97
Pneumothorax	7 / 109 (6.4)	6 / 111 (5.4)	1.09 (0.65 to 1.84)[b]	0.75
Atelectasis	87 / 109 (79.8)	95 / 111 (85.6)	0.82 (0.60 to 1.13)[b]	0.25
ICU length of stay, days	0.7 ± 1.0	0.7 ± 0.7		
Median (IQR)	0.3 (0.3 – 0.6)	0.4 (0.3 – 0.7)	-0.0 (-0.2 to 0.2)[a]	0.96

Table 3. Continued

	Fully automated ventilation (n = 109)	Conventional ventilation (n = 111)	Effect Estimate (95%–CI)	p value
ICU readmissions£££	2 / 109 (1.8)	4 / 111 (3.6)	0.67 (0.21 to 2.08)[b]	0.68
Mortality				
ICU	3 / 109 (2.8)d	0 / 111 (0.0)	---[e]	0.12[e]
30-day	3 / 109 (2.8)d	0 / 111 (0.0)	---[e]	0.08[e]
Percentage of breaths with SpO_2<85%¥,**	116 / 232,211 (0.0)	807 / 252,244 (0.3)	0.26 (0.22 to 0.31)[b]	< 0.01

** during the first three hours of ventilation or until extubation and for at least 30 consecutive seconds*

*** during the first three hours of ventilation or until extubation and reported according to the total number of breaths*

**** defined as any re-intubation within 48 hours after extubation and considering only patients who survived and did not undergo a re-sternotomy during this time*

***** time from ICU admission until first successful extubation*

£ time from stopping sedatives until successful extubation

££ time from ICU admission until ≥ 5 consecutives spontaneous breaths

£££ during the first 72 hours after ICU discharge

¥ when the measured SpO_2 had a quality index > 50%

a effect estimate is mean difference

b effect estimate is risk ratio

c effect estimate is hazard ratio

d One patient died the second day due to ventricular fibrillation as a result of cardiac ischaemia (postoperative coronary artery bypass graft failure); one patient died the second day as a result of an anastomotic rupture after aortic surgery; one patient died the fourth day due to a cardiac tamponade that occurred three days after extubation.

e since no event was observed in one group, the effect estimate was not calculated (infinite in the upper limit)

f 98.33% Bonferroni-corrected confidence intervals and p values adjusted for multiplicity according to Benjamini-Hochberg are:

Percentage of time in critical zone: -2.5 (-4.5 to -0.5); 0.003

Percentage of time in acceptable zone: -33.2 (-42.0 to -24.5); < 0.001

Percentage of time in optimal zone: 29.7 (20.4 to 39.1); < 0.001

Secondary outcomes

The time until the first spontaneous breathing effort was shorter (*Table 3*; Supplementary Figure S10) and the percentage of breaths with severe hypoxaemia was lower in the automated group. Duration of weaning and postoperative ventilation; the proportion of failed extubations; and developed postoperative pulmonary complications, ICU length of stay and readmission rates, and ICU and 30-day mortality were similar between automated and conventional ventilation groups (Supplementary Table S5).

Sensitivity and per-protocol analyses

Neither the per-protocol analysis (Supplementary Tables S5 and S6; Supplementary Figures S11 and S12) nor the sensitivity analysis (Supplementary Table S7) and post hoc analyses (Supplementary Table S8) altered the main findings. Differences in proportions of time spent in the ventilation zone between the automated group and the conventional group were similar in the two predefined subgroups (Supplementary Figures S13).

DISCUSSION

We found that fully automated ventilation increased the time patients were exposed to optimal, lung-protective settings, whilst reducing the risk of injurious ventilation. Automated ventilation was more likely to prevent severe hypoxaemia and accelerated the time until spontaneous breathing.

To the best of our knowledge, this is the largest study to date that compares fully automated closed-loop ventilation with conventional ventilation in patients after cardiac surgery receiving postoperative ventilation. The predefined primary outcome, which is comparable with a previous study of automated ventilation,[23] reflects both efficacy and safety of ventilation, although the study population had little pre-existing pulmonary pathology. The study was designed to minimise bias by using concealed allocation, collection of breath-by-breath data, a modified intention-to-treat analysis, and a pragmatic protocol. Of note, the protocol was strictly followed by a team of experienced and board-certified ICU nurses and physicians, resulting in high adherence to the protective ventilation strategy in the conventional group.

The present study has important differences compared with previous investigations of INTELLiVENT-ASV in patients after cardiac surgery.[23, 24] In the conventional group of the current study, VT and Pplat were lower than in the previous investigation (median VT <8 vs <10 ml kg^{-1} PBW; median Pplat <18 vs <21 cmH_2O),[23] and SpO_2 measurements were

closer to contemporary targets of oxygenation (median <98% vs 99%). In addition, the present study used a stricter definition for optimal ventilation. Indeed, in two previous studies,[23, 24] VT ≤10 ml kg^{-1} PBW was counted as optimal, whilst in the present study VT ≤8 ml kg^{-1} PBW counted as optimal. These differences explain why the reported proportion of time in the optimal zone was lower than in a previous study.[23]

Even though the absolute MD in the critical zone of –2.5% seems small in this study, a much larger difference can be expected in settings with less resources, less staff, and resource-poor training facilities. Notably, we found a large MD in the optimal zone of 27.2%. This study shows that INTELLiVENT-ASV results in ventilation with a lower VT, slightly higher PEEP, and a lower driving pressure compared with ventilation titrated by ICU nurses and doctors in an experienced specialist centre. Although evidence for benefit of ventilation with a lower VT and a lower driving pressure is most convincing in patients with acute respiratory distress syndrome (ARDS),[16, 28] there is increasing evidence for benefit of ventilation with a lower VT or a lower driving pressure in patients not having ARDS.[29–32] Even during intraoperative ventilation, use of a lower VT or a lower driving pressure has been found to be beneficial,[33–35] as was a high PEEP during postoperative ventilation in hypoxaemic cardiac surgery patients.[2]

Costs related to ICU patients are largely driven by costs pertaining to mechanically ventilated patients. Transforming the knowledge about protective ventilation into clinical practice is extremely challenging, but frequently time consuming and thus costly, which may result in inadequate and unsafe ventilatory support.[15, 36–38] Discrepancies between demand and supply are expected to become more common because of an ageing population and increasing severity of illness in patients.[39, 40] In addition, pandemics can put a huge strain on critical care resources, when systems have to struggle to provide high-quality care for a surge of critically ill patients in need of invasive ventilation. Fully automated ventilation modes could serve as a potential solution at minimal extra cost, whilst offering the potential to reduce the number of interactions with the ventilator by bedside caregivers.[20, 22, 25, 26] However, future studies are needed to determine the cost-effectiveness of fully automated ventilation for general ICU populations in resource-rich and resource-poor settings.

Our study has several limitations. Blinding was not possible because of the nature of intervention. The primary objective of this study was to determine the efficacy and safety of INTELLiVENT-ASV when compared with ventilation titrated by ICU nurses and doctors. The use of surrogate endpoints may not necessarily translate into better clinical outcomes. Future randomised clinical trials of this fully automated mode of ventilation need to explore patient-centred outcomes. Caution is needed when extrapolating

the results to other patient categories as the current study included a homogeneous cohort of only patients with minimal pre-existing pulmonary pathology who required postoperative mechanical ventilation for a relatively short period of time. Nonetheless, previous studies demonstrated that INTELLiVENT-ASV was safe and resulted in similar favourable improvements compared with conventional modes in critically ill patients with ARDS, COPD, or brain injury.[18, 19, 21, 22, 25–27] Also, as in previous studies, our study included haemodynamically stable patients.[20, 23, 24] Haemodynamic instability may interact unfavourably with automated ventilation software, because unstable patients with low cardiac output frequently have low $Etco_2$ and Spo_2 for haemodynamic reasons. In turn, these parameters may be 'misinterpreted' by the ventilator as a need for increase of minute ventilation and PEEP. Notably, in a previous study, the fully automated mode we tested performed similarly in fast-track cardiac surgery patients who were excluded in our study.[20] The plateau pressure was used as a surrogate measure for alveolar distending pressure to calculate the driving pressure and mechanical power. Whilst a direct measurement could have improved the accuracy of measurement, this was impractical in our study. Although the intention was to start the intervention as soon as the patient was admitted to the ICU, it was delayed until the results of the first blood gas analysis were available. However, the vast majority of patients commenced automated ventilation within 10 minutes after arrival in the ICU.

CONCLUSION

In this cohort of haemodynamic stable post-cardiac surgery patients receiving postoperative invasive ventilation by a team of well-trained and experienced ICU nurses and doctors, fully automated ventilation resulted in more likelihood of receiving lung-protective ventilation, fewer episodes of severe hypoxaemia, and more rapid return to spontaneous breathing. This study was not designed to evaluate other, more important patient-centred endpoints. Future studies should address whether fully automated ventilation is cost-effective in resource-rich and resource-poor settings.

REFERENCES

1. Lellouche F, Dionne S, Simard S, Bussieres J, Dagenais F. High tidal volumes in mechanically ventilated patients increase organ dysfunction after cardiac surgery. *Anesthesiology* 2012; **116**: 1072-82
2. Costa Leme A, Hajjar LA, Volpe MS, et al. Effect of Intensive vs Moderate Alveolar Recruitment Strategies Added to Lung-Protective Ventilation on Postoperative Pulmonary Complications: A Randomized Clinical Trial. *Jama* 2017; **317**: 1422-32
3. de Jonge E, Peelen L, Keijzers PJ, et al. Association between administered oxygen, arterial partial oxygen pressure and mortality in mechanically ventilated intensive care unit patients. *Critical care* 2008; **12**: R156
4. Eastwood G, Bellomo R, Bailey M, et al. Arterial oxygen tension and mortality in mechanically ventilated patients. *Intensive care medicine* 2012; **38**: 91-8
5. Girardis M, Busani S, Damiani E, et al. Effect of Conservative vs Conventional Oxygen Therapy on Mortality Among Patients in an Intensive Care Unit: The Oxygen-ICU Randomized Clinical Trial. *Jama* 2016; **316**: 1583-9
6. Helmerhorst HJ, Arts DL, Schultz MJ, et al. Metrics of Arterial Hyperoxia and Associated Outcomes in Critical Care. *Critical care medicine* 2017; **45**: 187-95
7. Asfar P, Schortgen F, Boisrame-Helms J, et al. Hyperoxia and hypertonic saline in patients with septic shock (HYPERS2S): a two-by-two factorial, multicentre, randomised, clinical trial. *The Lancet Respiratory medicine* 2017; **5**: 180-90
8. Chu DK, Kim LH, Young PJ, et al. Mortality and morbidity in acutely ill adults treated with liberal versus conservative oxygen therapy (IOTA): a systematic review and meta-analysis. *Lancet* 2018; **391**: 1693-705
9. Mackle D, Bellomo R, Bailey M, et al. Conservative Oxygen Therapy during Mechanical Ventilation in the ICU. *The New England journal of medicine* 2019
10. Wenstedt EFE, De Bie Dekker AJR, Roos AN, et al. Current practice of closed-loop mechanical ventilation modes on intensive care units - A nationwide survey in the Netherlands. *Netherlands Journal of Medicine* 2017; **75**: 145-50
11. Chatburn RL, Mireles-Cabodevila E. Closed-loop control of mechanical ventilation: description and classification of targeting schemes. *Respiratory care* 2011; **56**: 85-102
12. McConville JF, Kress JP. Weaning patients from the ventilator. *The New England journal of medicine* 2012; **367**: 2233-9
13. Burns KE, Lellouche F, Nisenbaum R, Lessard MR, Friedrich JO. Automated weaning and SBT systems versus non-automated weaning strategies for weaning time in invasively ventilated critically ill adults. *The Cochrane database of systematic reviews* 2014: CD008638
14. Rose L, Schultz MJ, Cardwell CR, Jouvet P, McAuley DF, Blackwood B. Automated versus non-automated weaning for reducing the duration of mechanical ventilation for critically ill adults and children: a cochrane systematic review and meta-analysis. *Critical care* 2015; **19**: 48
15. Dellaca RL, Veneroni C, Farre R. Trends in mechanical ventilation: are we ventilating our patients in the best possible way? *Breathe* 2017; **13**: 84-98

16. Brower RG, Matthay MA, Morris A, Schoenfeld D, Thompson BT, Wheeler A. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *The New England journal of medicine* 2000; **342**: 1301-8
17. Lellouche F, Brochard L. Advanced closed loops during mechanical ventilation (PAV, NAVA, ASV, SmartCare). *Best practice & research Clinical anaesthesiology* 2009; **23**: 81-93
18. Arnal JM, Wysocki M, Novotni D, et al. Safety and efficacy of a fully closed-loop control ventilation (IntelliVent-ASV) in sedated ICU patients with acute respiratory failure: a prospective randomized crossover study. *Intensive care medicine* 2012; **38**: 781-7
19. Clavieras N, Wysocki M, Coisel Y, et al. Prospective randomized crossover study of a new closed-loop control system versus pressure support during weaning from mechanical ventilation. *Anesthesiology* 2013; **119**: 631-41
20. Beijers AJ, Roos AN, Bindels AJ. Fully automated closed-loop ventilation is safe and effective in post-cardiac surgery patients. *Intensive care medicine* 2014; **40**: 752-3
21. Abutbul A SS, Zbedat W, Linton DM, van Heerden PV. A prospective comparison of the efficacy and safety of fully closed-loop control ventilation (Intellivent-ASV) with conventional ASV and SIMV modes. *S Afr J Crit Care* 2014; **30**: 28-32
22. Bialais E, Wittebole X, Vignaux L, et al. Closed-loop ventilation mode (IntelliVent(R)-ASV) in intensive care unit: a randomized trial. *Minerva anestesiologica* 2016; **82**: 657-68
23. Lellouche F, Bouchard PA, Simard S, L'Her E, Wysocki M. Evaluation of fully automated ventilation: a randomized controlled study in post-cardiac surgery patients. *Intensive care medicine* 2013; **39**: 463-71
24. Fot EV, Izotova NN, Yudina AS, Smetkin AA, Kuzkov VV, Kirov MY. Automated Weaning from Mechanical Ventilation after Off-Pump Coronary Artery Bypass Grafting. *Frontiers in medicine* 2017; **4**
25. Anan'ev EP, Polupan AA, Matskovskiy IV, et al. Use of the IntelliVent-ASV mode for maintaining the target EtCO2 range in patients with severe TBI. *NN Burdenko's journal of neurosurgery* 2017; **81**: 63-8
26. Arnal JM, Garnero A, Novotni D, et al. Closed loop ventilation mode in Intensive Care Unit: A randomized controlled clinical trial comparing the numbers of manual ventilator setting changes. *Minerva anestesiologica* 2018; **84**: 58-67
27. Arnal JM, Saoli M, Garnero A. Airway and transpulmonary driving pressures and mechanical powers selected by INTELLiVENT-ASV in passive, mechanically ventilated ICU patients. *Heart & Lung* 2019; **-**: 1-8
28. Amato MB, Meade MO, Slutsky AS, et al. Driving pressure and survival in the acute respiratory distress syndrome. *The New England journal of medicine* 2015; **372**: 747-55
29. Serpa Neto A, Simonis FD, Barbas CS, et al. Association between tidal volume size, duration of ventilation, and sedation needs in patients without acute respiratory distress syndrome: an individual patient data meta-analysis. *Intensive care medicine* 2014; **40**: 950-7

30. Serpa Neto A, Simonis FD, Barbas CSV, et al. Lung-Protective Ventilation With Low Tidal Volumes and the Occurrence of Pulmonary Complications in Patients Without Acute Respiratory Distress Syndrome: A Systematic Review and Individual Patient Data Analysis. *Read Online: Critical Care Medicine | Society of Critical Care Medicine* 2015; **43**: 2155-63
31. Serpa Neto A, Deliberato RO, Johnson AEW, et al. Mechanical power of ventilation is associated with mortality in critically ill patients: an analysis of patients in two observational cohorts. *Intensive care medicine* 2018; **44**: 1914-22
32. Lanspa MJ, Peltan ID, Jacobs JR, et al. Driving pressure is not associated with mortality in mechanically ventilated patients without ARDS. *Critical care* 2019; **23**: 424
33. Serpa Neto A, Hemmes SNT, Barbas CSV, et al. Protective versus Conventional Ventilation for Surgery: A Systematic Review and Individual Patient Data Meta-analysis. *Anesthesiology: The Journal of the American Society of Anesthesiologists* 2015; **123**: 66-78
34. Serpa Neto A, Hemmes SNT, Barbas CSV, et al. Association between driving pressure and development of postoperative pulmonary complications in patients undergoing mechanical ventilation for general anaesthesia: a meta-analysis of individual patient data. *The Lancet Respiratory Medicine* 2016; **4**: 272-80
35. Young CC, Harris EM, Vacchiano C, et al. Lung-protective ventilation for the surgical patient: international expert panel-based consensus recommendations. *British journal of anaesthesia* 2019; **123**: 898-913
36. Rubenfeld GD, Cooper C, Carter G, Thompson BT, Hudson LD. Barriers to providing lung-protective ventilation to patients with acute lung injury. *Critical care medicine* 2004; **32**: 1289-93
37. Young MP, Manning HL, Wilson DL, et al. Ventilation of patients with acute lung injury and acute respiratory distress syndrome: has new evidence changed clinical practice? *Critical care medicine* 2004; **32**: 1260-5
38. Mikkelsen ME, Dedhiya PM, Kalhan R, Gallop RJ, Lanken PN, Fuchs BD. Potential reasons why physicians underuse lung-protective ventilation: a retrospective cohort study using physician documentation. *Respiratory care* 2008; **53**: 455-61
39. Beet C, Benoit D, Bion J. Intensive care medicine in 2050: preventing harm. *Intensive care medicine* 2019; **45**: 505-7
40. Harlem Brundtland GHE, Sy EA, Dzau V, et al. A world at risk: annual report on global preparedness for health emergencies. *World Health Organization* 2019

SUPPLEMENTAL DIGITAL CONTENT

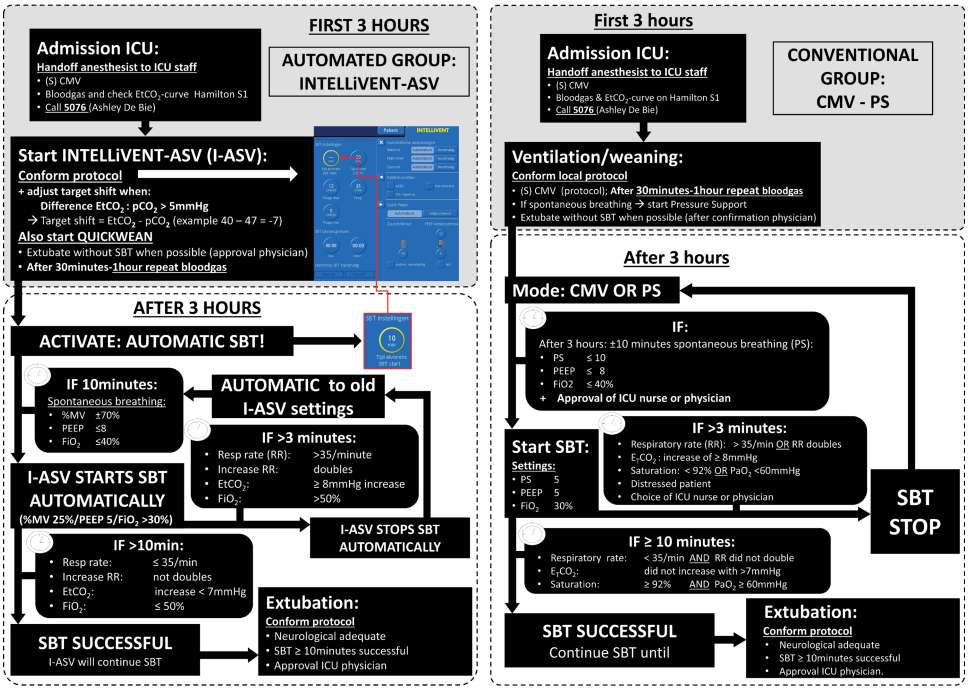


Figure S1. Standard operation procedures used at the bedside describing the ventilation strategy in the both groups

Abbreviations

(S) CMV: Continuous mandatory ventilation; $etCO_2$: end-tidal carbon dioxide; FIO_2: inspired fraction of oxygen; ICU: intensive care unit; PaO_2: partial pressure of arterial oxygen; PEEP: positive end-expiratory pressure; PS: Pressure support; RR: respiratory rate; SBT: spontaneous breathing trial; SpO_2: oxygen pulse oximetry.

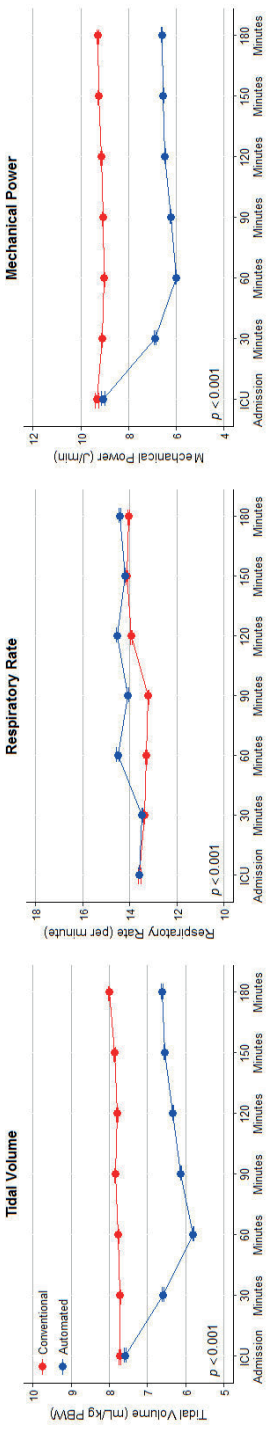


Figure S2. Tidal volume, respiratory rate and mechanical power every 30 minutes until 180 minutes

p value from a mixed-effect longitudinal models with random intercepts for patients; time was treated as a continuous variable

Figure S3. FIO_2, $etCO_2$ and SpO_2 every 30 minutes until 180 minutes.

FIO2: inspired fraction of oxygen; $etCO_2$: end-tidal carbon dioxide; SpO_2: oxygen pulse oximetry; p value from a mixed-effect longitudinal models with random intercepts for patients; time was treated as a continuous variable

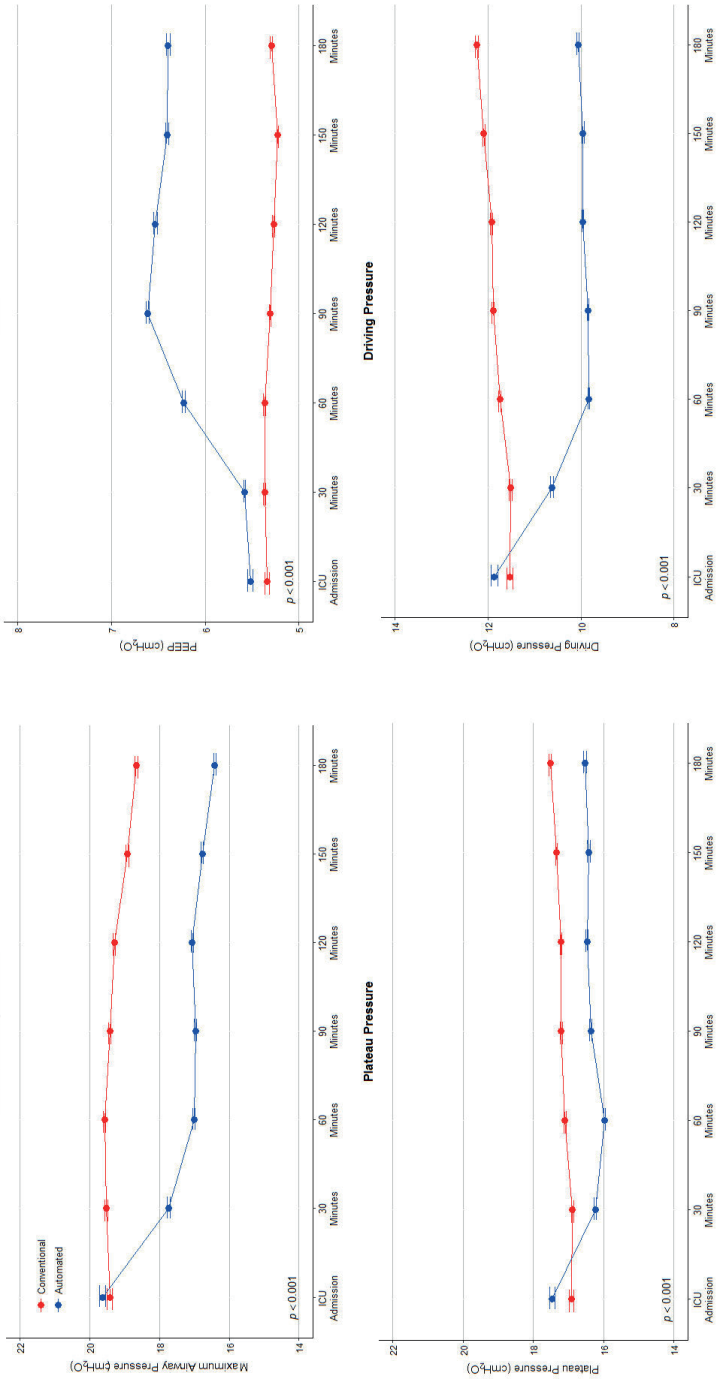


Figure S4. Maximum airway pressure, plateau pressure, driving pressure and PEEP every 30 minutes until 180 minutes

PEEP: positive end-expiratory pressure. p value from a mixed-effect longitudinal models with random intercepts for patients; time was treated as a continuous variable

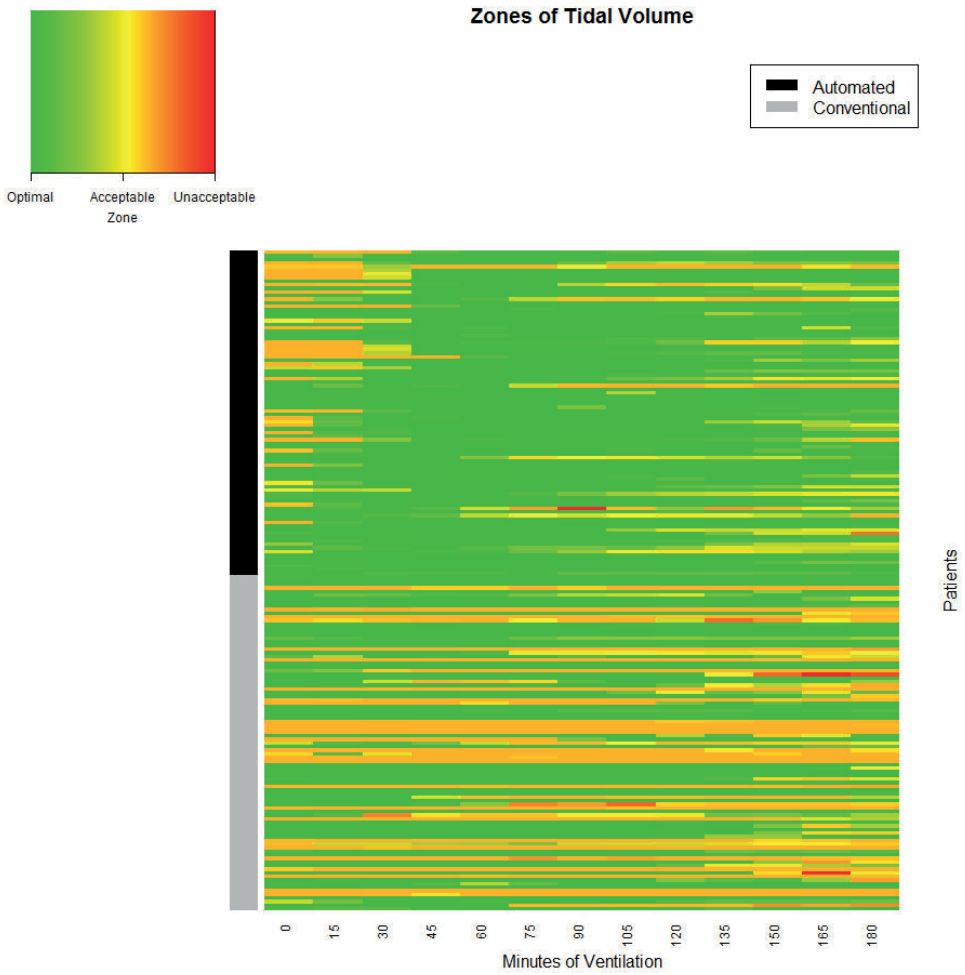


Figure S5. Heat map showing the ventilation zones for tidal volume every 15 minutes after randomisation

For the heat map construction, each breath was assigned to the corresponding ventilation zone and was given a numeric value as follows: (1) for a breath in the optimal zone, (2) for a breath in the acceptable zone, and (3) for a breath in the unacceptable zone. Then, all value of the breaths recorded within every 15 minutes were summarised using the mean of these values. The default colour gradient sets the lowest mean value to green (Optimal), the highest value to a bright red (Unacceptable), and mid-range values to yellow (Acceptable), with a corresponding transition between these extremes

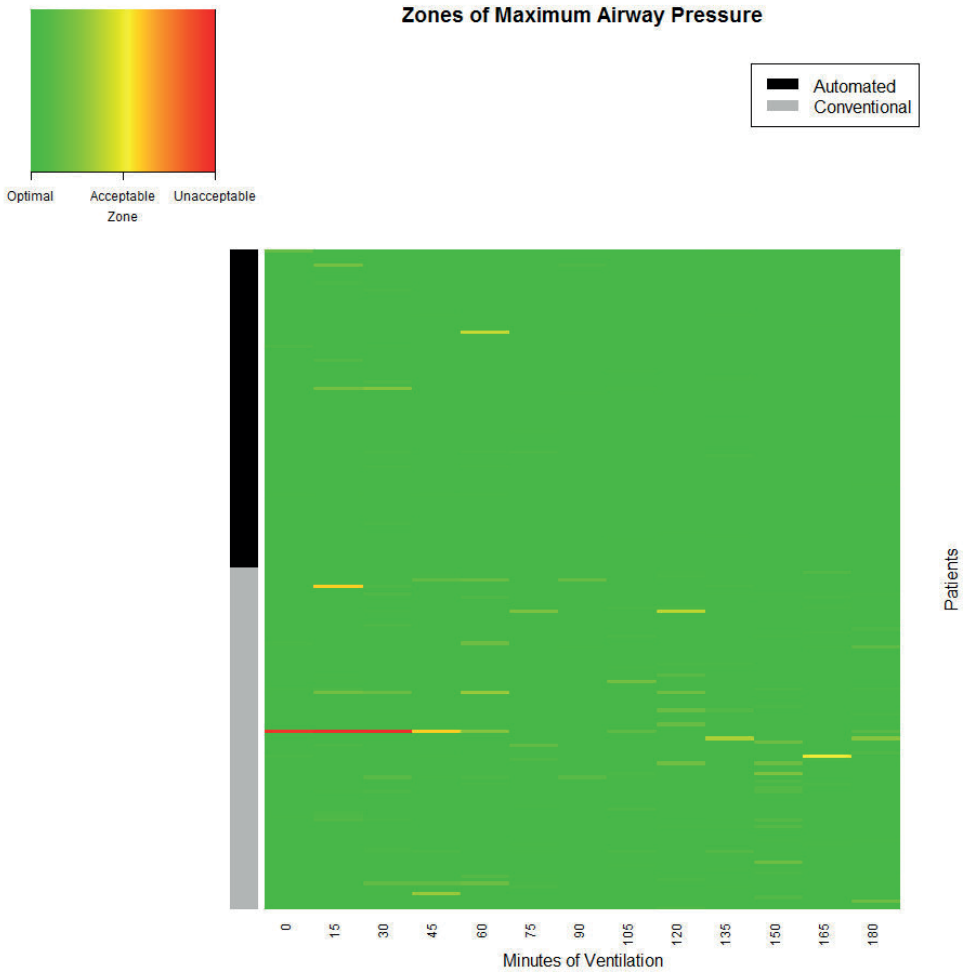


Figure S6. Heat map showing the ventilation zones for maximum airway pressure every 15 minutes after randomisation

For the heat map construction, each breath was assigned to the corresponding ventilation zone and was given a numeric value as follows: (1) for a breath in the optimal zone, (2) for a breath in the acceptable zone, and (3) for a breath in the unacceptable zone. Then, all value of the breaths recorded within every 15 minutes were summarised using the mean of these values. The default colour gradient sets the lowest mean value to green (Optimal), the highest value to a bright red (Unacceptable), and mid-range values to yellow (Acceptable), with a corresponding transition between these extremes.

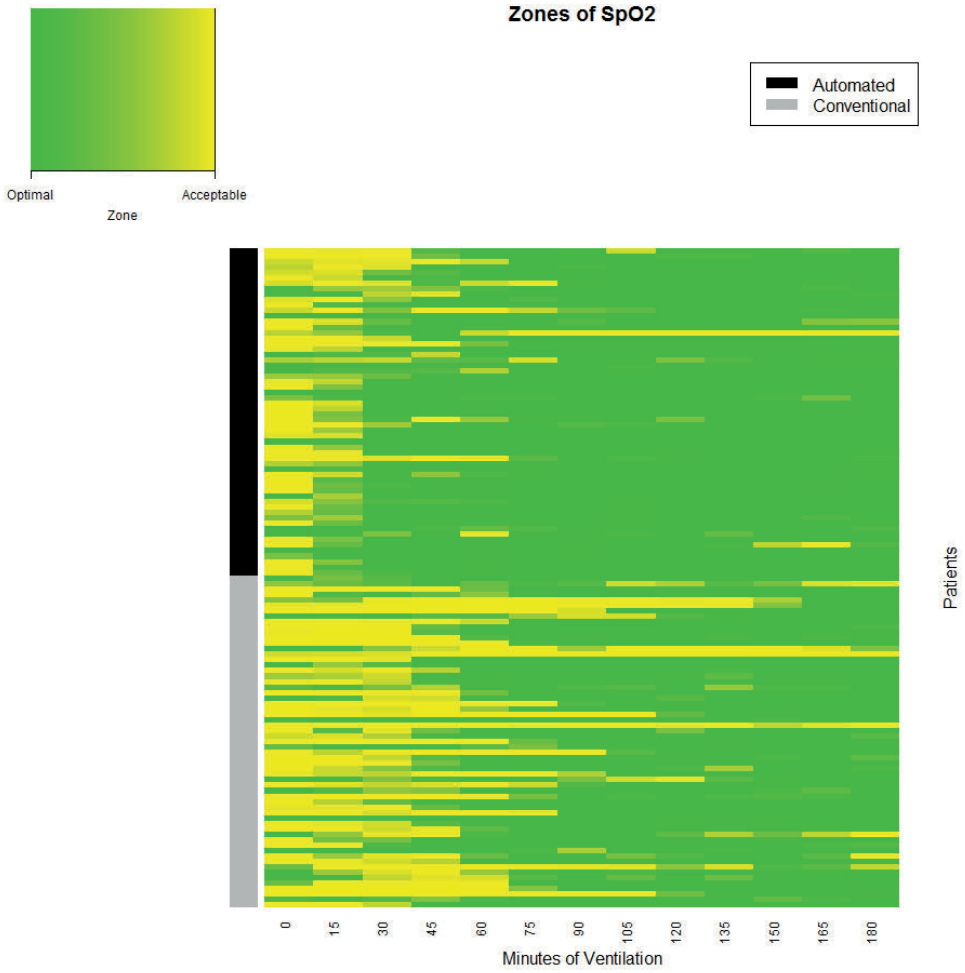


Figure S7. Heat map showing the ventilation zones for SpO_2 every 15 minutes after randomisation

For the heat map construction, each breath was assigned to the corresponding ventilation zone and was given a numeric value as follows: (1) for a breath in the optimal zone, (2) for a breath in the acceptable zone, and (3) for a breath in the unacceptable zone. Then, all value of the breaths recorded within every 15 minutes were summarised using the mean of these values. The default colour gradient sets the lowest mean value to green (Optimal), the highest value to a bright red (Unacceptable), and mid-range values to yellow (Acceptable), with a corresponding transition between these extremes.

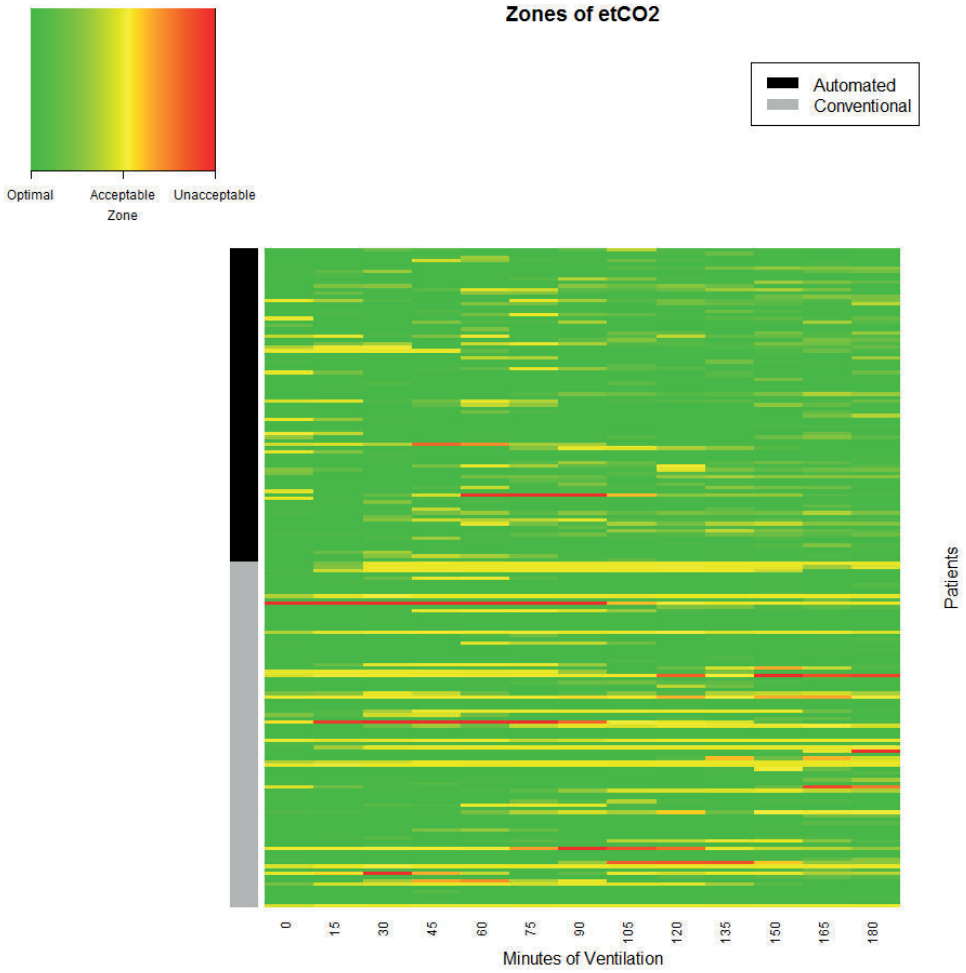


Figure S8. Heat map showing the ventilation zones for $EtCO_2$ every 15 minutes after randomisation
For the heat map construction, each breath was assigned to the corresponding ventilation zone and was given a numeric value as follows: (1) for a breath in the optimal zone, (2) for a breath in the acceptable zone, and (3) for a breath in the unacceptable zone. Then, all value of the breaths recorded within every 15 minutes were summarised using the mean of these values. The default colour gradient sets the lowest mean value to green (Optimal), the highest value to a bright red (Unacceptable), and mid-range values to yellow (Acceptable), with a corresponding transition between these extremes.

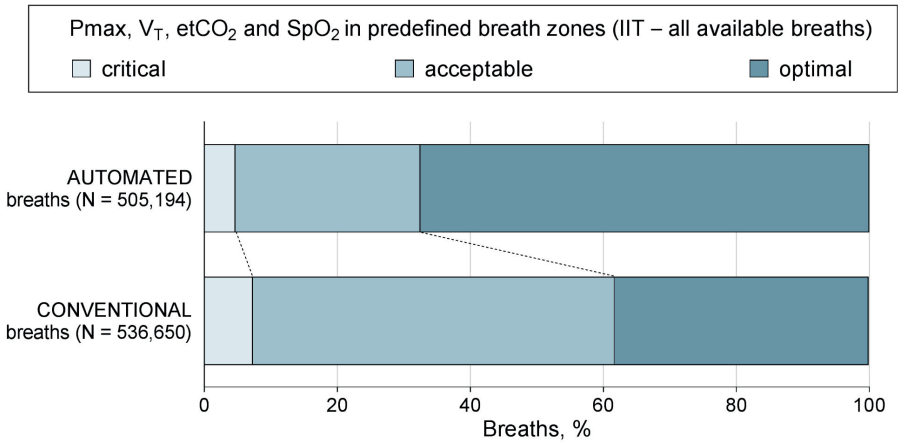


Figure S9. Percentage of breaths in pre-defined zones during the complete ventilation time in the intention-to-treat population

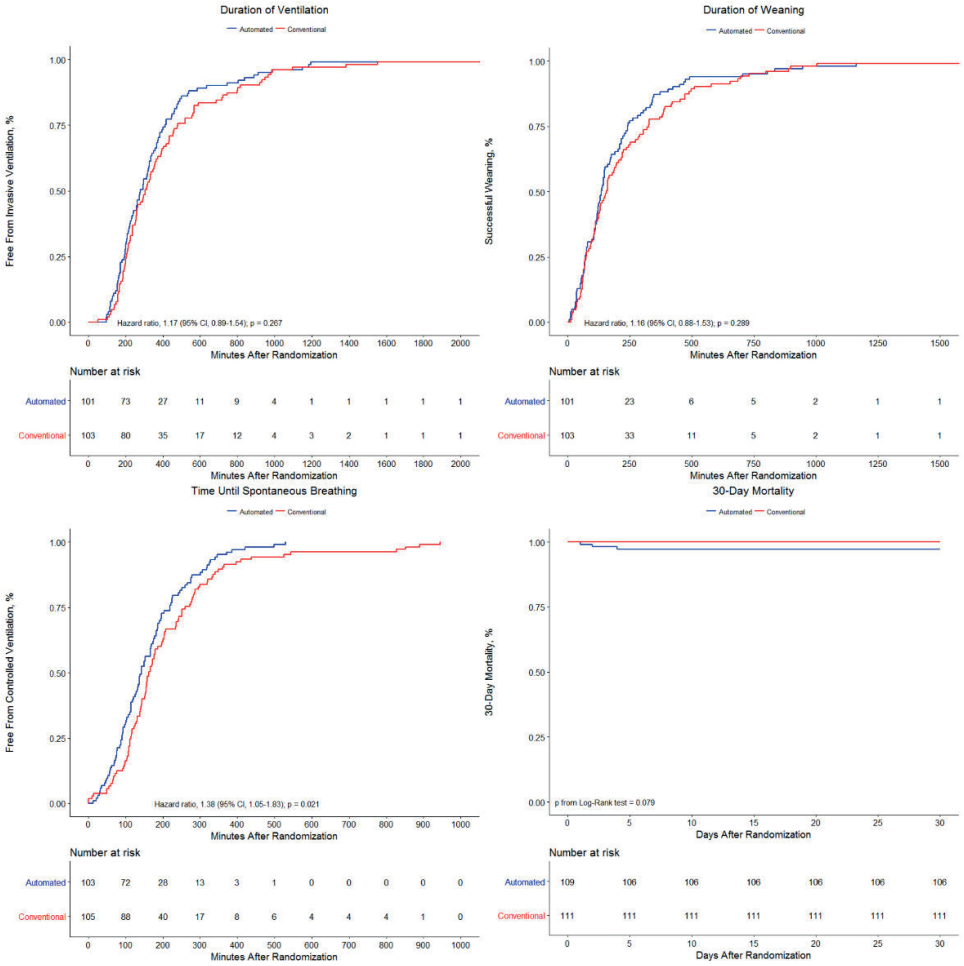


Figure S10. Kaplan–Meier Estimates for Patients in the Automated and Conventional Groups

A, *Median (IQR) observation period for the duration free from invasive ventilation was 279 minutes (195 to 412) for the automated group and 304 minutes (204 to 479) for the conventional group; p value for the Schoenfeld residuals was 0.570.*

B, *Median (IQR) observation period for the successful weaning was 136 minutes (73 to 241) for the automated group and 157 minutes (75 to 372) for the conventional group; p value for the Schoenfeld residuals was 0.927.*

C, *Median (IQR) observation period for the duration free from controlled ventilation was 142 minutes (90 to 219) for the automated group and 162 minutes (114 to 260) for the conventional group; p value for the Schoenfeld residuals was 0.793.*

D, *Median observation time for survival was not computed for 30-day mortality because the minimum value observed is 0.97; p value for the Schoenfeld residuals was 0.999*

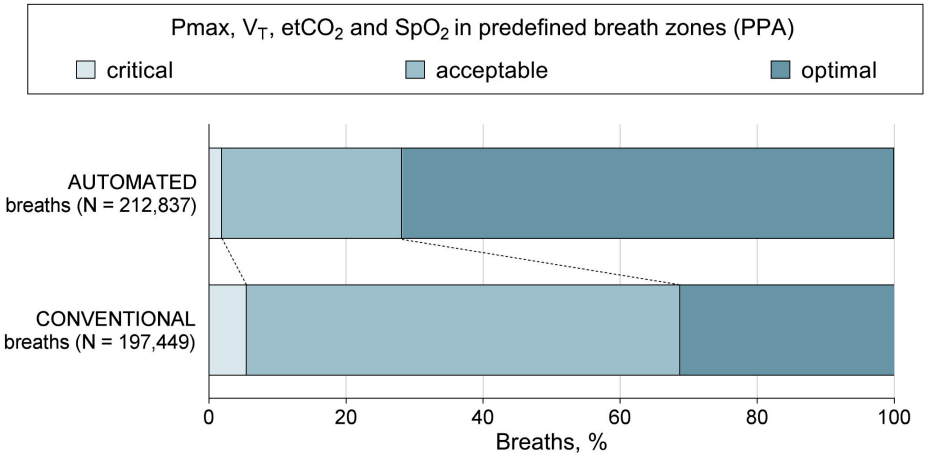


Figure S11. Percentage of breaths in pre-defined zones during the first three hours of ventilation or until extubation in the per-protocol population

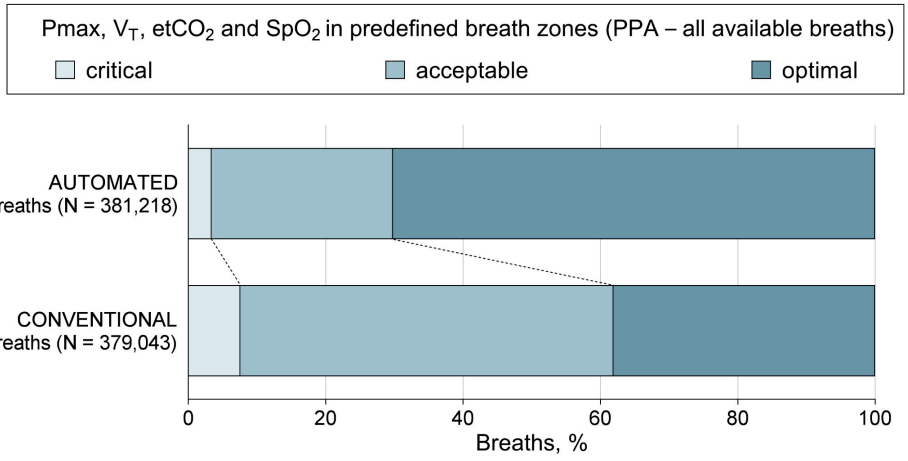


Figure S12. Percentage of breaths in pre-defined zones during the complete ventilation time in the *per*-protocol population

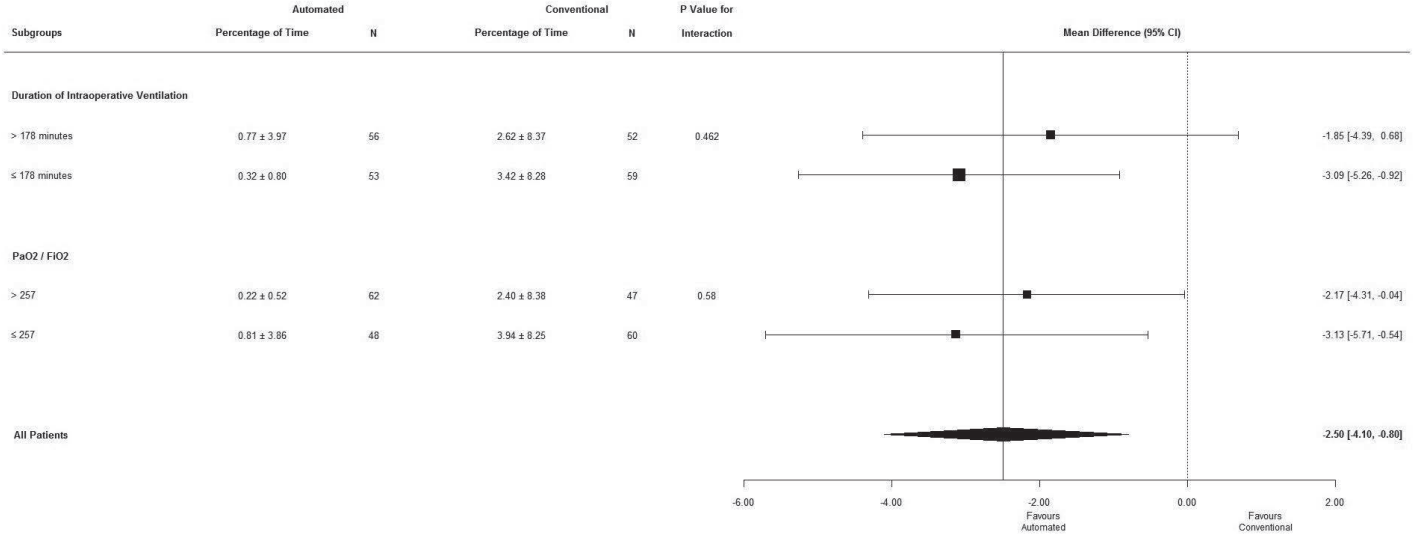


Figure S13. Effects of the automated vs conventional strategy of ventilation on proportion of time spent in critical ventilation zone according to pre-specified subgroups.

The size of the markers is proportional to the number of patients per group. Vertical solid line represents the overall effect and vertical dashed line represents the reference.

Table S1. Provided dosages of intravenous sedatives and analgesia during surgery, ICU admission, and both combined

	Fully automated ventilation (*n* = 109)	**Conventional ventilation (*n* = 111)**	***p value***
During surgery			
Etomidate (mg)	50 (50 – 70)	50 (50 – 50)	0.29
Rocuronium (mg)	200 (200 – 200)	200 (200 – 200)	0.03
Propofol (mg)	1437 (1135 – 1760)	1404 (1130 – 1782)	0.33
Midazolam (mg)	2.5 (0 – 5)	5 (0 – 5)	0.64
Opiates			
Morphine (mg)	25 (25 – 25)	25 (25 – 25)	0.31
Alfentanil (mg)	1233 (955 – 1533)	1189 (1015 – 1545)	0.98
Sufentanil (mcg)	0 (0 – 0)	0 (0 – 0)	0.31
During ICU admission			
Etomidate (mg)	-	-	-
Rocuronium (mg)	-	-	-
Propofol (mg)	178 (84 – 301)	178 (95 – 305)	0.71
Midazolam (mg)	0 (0 – 0)	0 (0 – 0)	0.33
Opiates			
Morphine (mg)	0 (0-2)	0 (0 – 2)	1.00
Alfentanil (mg)	-	-	-
Sufentanil (mcg)	-	-	-
Total amount			
Etomidate (mg)	50 (50 – 70)	50 (50 – 70)	0.29
Rocuronium (mg)	200 (200 – 200)	200 (200 – 200)	0.03
Propofol (mg)	1659 (1262 – 2120)	1602 (1371 – 2052)	0.87
Midazolam (mg)	1 (0 – 5)	4 (0 – 5)	0.26
Opiates			
Morphine (mg)	25 (25 – 27)	25 (25 – 27)	0.94
Alfentanil (mg)	1233 (955 – 1533)	1189 (1015 – 1545)	0.98
Sufentanil (mcg)	0 (0 – 0)	0 (0 – 0)	0.31

Table S2. Ventilatory variables over time

	At ICU admission*			After 30 minutes	
	Automated	**Conventional**	***p* value**	**Automated**	**Conventional**
Number of patients	109	111	---	109	111
Number of breaths	7,130	7,293	---	36,065	36,529
Tidal volume, mL kg^{-1} PBW	7.6 ± 1.1	7.7 ± 1.1	< 0.001	6.6 ± 1.5	7.7 ± 1.3
PEEP, cmH_2O	5.5 ± 1.2	5.3 ± 1.0	< 0.001	5.6 ± 1.1	5.4 ± 1.1
Maximum airway pressure, cmH_2O	19.6 ± 3.4	19.4 ± 3.4	< 0.001	17.7 ± 3.8	19.5 ± 3.5
Plateau pressure, cmH_2O*	17.5 ± 2.8	16.9 ± 2.8	< 0.001	16.2 ± 2.8	16.9 ± 2.6
Driving pressure, cmH_2O*	11.9 ± 2.6	11.5 ± 2.6	< 0.001	10.6 ± 2.6	11.5 ± 2.3
Mechanical Power (J min^{-1})*	9.1 ± 2.6	9.3 ± 2.5	< 0.001	6.9 ± 2.6	9.1 ± 2.3
Respiratory rate, bpm	13.6 ± 2.7	13.6 ± 2.4	0.892	13.4 ± 2.1	13.4 ± 1.9
Static compliance, mL cmH_2O^{-1}	43.5 ± 10.7	45.5 ± 9.2	< 0.001	41.6 ± 10.6	45.4 ± 9.1
FIO_2, %	50.9 ± 7.1	52.1 ± 6.7	< 0.001	44.3 ± 10.2	51.8 ± 7.3
SpO_2, %	97.9 ± 2.2	97.4 ± 2.9	< 0.001	97.2 ± 2.5	97.5 ± 3.1
$EtCO_2$, kPa	4.6 ± 0.6	4.6 ± 0.7	< 0.001	4.9 ± 0.7	4.5 ± 0.7
	At ICU admission*			**After 120 minutes**	
	Automated	**Conventional**	***p* value**	**Automated**	**Conventional**
Number of patients	109	111	---	107	108
Number of breaths	7,130	7,293	---	43,202	42,999
Tidal volume, mL kg^{-1} PBW	7.6 ± 1.1	7.7 ± 1.1	< 0.001	6.3 ± 1.9	7.8 ± 1.7
PEEP, cmH_2O	5.5 ± 1.2	5.3 ± 1.0	< 0.001	6.5 ± 2.0	5.3 ± 1.1
Maximum airway pressure, cmH_2O	19.6 ± 3.4	19.4 ± 3.4	< 0.001	17.0 ± 3.3	19.3 ± 3.4
Plateau pressure, cmH_2O	17.5 ± 2.8	16.9 ± 2.8	< 0.001	16.5 ± 2.9	17.2 ± 2.4
Driving pressure, cmH_2O	11.9 ± 2.6	11.5 ± 2.6	< 0.001	10.0 ± 2.0	11.9 ± 2.2
Mechanical Power (J min^{-1})	9.1 ± 2.6	9.3 ± 2.5	< 0.001	6.5 ± 2.4	9.1 ± 2.2
Respiratory rate, bpm	13.6 ± 2.7	13.6 ± 2.4	0.892	14.5 ± 3.8	13.9 ± 3.5
Static compliance, mL cmH_2O^{-1}	43.5 ± 10.7	45.5 ± 9.2	< 0.001	40.4 ± 12.4	43.5 ± 10.1
FIO_2, %	50.9 ± 7.1	52.1 ± 6.7	< 0.001	32.1 ± 6.4	42.2 ± 5.5
SpO_2, %	97.9 ± 2.2	97.4 ± 2.9	< 0.001	96.6 ± 2.2	97.8 ± 2.5
$EtCO_2$, kPa	4.6 ± 0.6	4.6 ± 0.7	< 0.001	5.6 ± 0.5	4.6 ± 0.9

Abbreviations: PEEP: positive end-expiratory pressure; PBW: predicted body weight; FIO_2: inspired fraction of oxygen; SpO_2: oxygen pulse oximetry; $etCO_2$: end-tidal carbon dioxide

** Considering the first 5 minutes after ICU admission*

	After 60 minutes			After 90 minutes		
***p* value**	**Automated**	**Conventional**	***p* value**	**Automated**	**Conventional**	***p* value**
---	109	111	---	107	109	---
---	44,802	43,172	---	43,900	42,222	---
< 0.001	5.8 ± 1.4	7.8 ± 1.3	< 0.001	6.1 ± 2.0	7.8 ± 1.4	< 0.001
< 0.001	6.2 ± 1.6	5.4 ± 1.1	< 0.001	6.6 ± 2.0	5.3 ± 1.0	< 0.001
< 0.001	17.0 ± 3.2	19.6 ± 3.4	< 0.001	17.0 ± 3.0	19.4 ± 3.3	< 0.001
< 0.001	16.0 ± 2.5	17.1 ± 2.6	< 0.001	16.4 ± 2.7	17.2 ± 2.5	< 0.001
< 0.001	9.8 ± 1.9	11.7 ± 2.1	< 0.001	9.8 ± 1.9	11.9 ± 2.1	< 0.001
< 0.001	6.0 ± 2.3	9.0 ± 2.2	< 0.001	6.2 ± 2.1	9.1 ± 2.7	< 0.001
< 0.001	14.5 ± 5.0	13.3 ± 2.5	< 0.001	14.1 ± 2.7	13.2 ± 2.0	< 0.001
< 0.001	39.6 ± 10.7	44.4 ± 8.4	< 0.001	39.9 ± 11.7	43.8 ± 9.1	< 0.001
< 0.001	35.7 ± 10.1	44.4 ± 6.9	< 0.001	33.0 ± 7.2	44.1 ± 6.5	< 0.001
< 0.001	96.0 ± 2.6	98.1 ± 2.5	< 0.001	96.4 ± 2.1	98.0 ± 2.1	< 0.001
< 0.001	5.5 ± 0.7	4.5 ± 0.8	< 0.001	5.6 ± 0.6	4.6 ± 0.8	< 0.001
	After 150 minutes			**After 180 minutes**		
***p* value**	**Automated**	**Conventional**	***p* value**	**Automated**	**Conventional**	***p* value**
---	99	106	---	95	102	---
---	39,717	41,759	---	37,015	38,283	---
< 0.001	6.5 ± 2.0	7.8 ± 2.0	< 0.001	6.6 ± 2.1	8.0 ± 2.2	< 0.001
< 0.001	6.4 ± 2.0	5.2 ± 1.1	< 0.001	6.4 ± 2.0	5.3 ± 1.2	< 0.001
< 0.001	16.8 ± 3.3	18.9 ± 3.8	< 0.001	16.4 ± 3.2	18.7 ± 4.0	< 0.001
< 0.001	16.4 ± 3.0	17.3 ± 2.4	< 0.001	16.5 ± 3.0	17.5 ± 2.5	< 0.001
< 0.001	9.9 ± 2.1	12.1 ± 2.4	< 0.001	10.1 ± 2.2	12.2 ± 2.4	< 0.001
< 0.001	6.5 ± 2.4	9.3 ± 2.5	< 0.001	6.6 ± 2.4	9.3 ± 2.5	< 0.001
< 0.001	14.2 ± 3.2	14.1 ± 3.5	< 0.001	14.4 ± 3.3	14.0 ± 4.0	< 0.001
< 0.001	41.0 ± 14.6	42.2 ± 9.4	< 0.001	40.7 ± 16.6	41.9 ± 10.3	< 0.001
< 0.001	32.9 ± 9.1	41.3 ± 5.2	< 0.001	33.3 ± 9.0	41.9 ± 5.0	< 0.001
< 0.001	96.7 ± 2.2	97.5 ± 3.0	< 0.001	96.6 ± 2.2	97.6 ± 2.9	< 0.001
< 0.001	5.6 ± 0.5	4.7 ± 0.9	< 0.001	5.6 ± 0.6	4.7 ± 0.9	< 0.001

Table S3. Arterial blood gas analyses

	First			Second	
	Automated	**Conventional**	***p* value**	**Automated**	**Conventional**
N	**107**	**110**	**---**	**87**	**92**
Mean time from admission, minutes	20.8 ± 16.7	20.7 ± 13.8	0.948	105.1 ± 42.9	112.6 ± 69.2
$EtCO_2$ from ventilator, kPa	4.8 ± 0.6	4.5 ± 0.8	0.128	5.9 ± 1.4	4.7 ± 0.8
$PaCO_2$, kPa	5.3 ± 0.8	5.1 ± 0.9	0.204	5.8 ± 0.6	5.3 ± 1.4
SpO_2 from ventilator, %	97.6 ± 2.6	98.4 ± 2.0	0.240	96.5 ± 2.2	98.2 + 2.1
SaO_2, %	97.7 ± 1.8	98.3 ± 1.9	0.032	96.2 ± 1.9	98.2 ± 5.9
PaO_2, kPa	16.7 ± 4.8	19.0 ± 5.4	< 0.001	13.1 ± 2.6	16.9 ± 4.7
pH	7.34 ± 0.06	7.34 ± 0.06	0.729	7.30 ± 0.05	7.34 ± 0.06
Bicarbonate, mEq L^{-1}	20.7 ± 2.3	20.2 ± 1.9	0.089	21.2 ± 2.1	20.4 ± 2.2
Lactate, mmol L^{-1}	1.7 ± 0.7	1.6 ± 0.6	0.281	1.6 ± 1.0	1.4 ± 0.7

SpO_2: oxygen pulse oximetry; $EtCO_2$: end-tidal carbon dioxide; $PaCO_2$: partial pressure of carbon dioxide; PaO_2: partial pressure of oxygen; SaO_2: arterial saturation of oxygen

	Third			Fourth		
***p* value**	**Automated**	**Conventional**	***p* value**	**Automated**	**Conventional**	***p* value**
---	**40**	**42**	**---**	**18**	**19**	**---**
0.382	293.3 ± 127.6	295.2 ± 143.1	0.950	543.0 ± 216.6	537.5 ± 153.5	0.930
< 0.001	5.6 ± 0.5	5.1 ± 0.6	0.068	5.3 ± 0.7	4.9 ± 0.7	0.358
0.001	5.7 ± 0.6	5.3 ± 0.6	0.041	5.3 ± 0.7	5.2 ± 0.9	0.562
0.015	96.3 ± 2.4	97.5 ± 2.1	0.247	96.2 ± 3.3	97.3 ± 1.5	0.576
0.003	96.4 ± 1.6	97.1 ± 2.0	0.110	96.7 ± 1.6	97.5 ± 1.5	0.141
< 0.001	13.3 ± 4.0	15.2 ± 4.0	0.037	13.2 ± 2.0	14.6 ± 3.4	0.134
< 0.001	7.31 ± 0.05	7.33 ± 0.05	0.119	7.32 ± 0.06	7.35 ± 0.06	0.143
0.011	21.2 ± 2.6	20.7 ± 2.5	0.358	20.2 ± 3.3	20.5 ± 3.4	0.761
0.303	1.7 ± 0.7	1.8 ± 1.1	0.705	2.9 ± 2.1	1.9 ± 1.4	0.124

Table S4. Co-primary and secondary outcomes for all recorded breaths and time in the intention-to-treat population

	Fully automated ventilation (*n* = 109)	**Conventional ventilation (*n* = 111)**	**Effect Estimate (95% CI)**	***p value***
Primary outcome				
Percentage of time in the optimal zone*	46.8 ± 26.0	25.8 ± 24.0	21.0 (14.3 to 27.7)[a]	< 0.001
Median (IQR)	47.2 (26.3 – 66.6)	20.4 (2.6 – 44.5)		
Percentage of time in the acceptable zone*	16.1 ± 15.0	39.8 ± 27.8	-23.7 (-29.6 to -17.7)[a]	< 0.001
Median (IQR)	12.2 (5.3 – 21.8)	32.6 (15.9 – 62.8)		
Percentage of time in the critical zone*	1.5 ± 4.7	3.6 ± 8.1	-2.1 (-3.8 to -0.3)[a]	0.020
Median (IQR)	0.0 (0.0 – 0.5)	0.7 (0.0 – 1.9)		
Secondary outcomes				
Percentage of breaths in the optimal zone**	340,889 / 505,194 (67.5)	205,232 / 536,650 (38.2)	1.88 (1.87 to 1.89)[b]	< 0.001
Percentage of breaths in the acceptable zone**	140,595 / 505,194 (27.8)	292,489 / 536,650 (54.4)	0.54 (0.53 to 0.54)[b]	< 0.001
Percentage of breaths in the critical zone**	23,710 / 505,194 (4.7)	38,929 / 536,650 (7.3)	0.76 (0.76 to 0.77)[b]	< 0.001
Percentage of breaths with SpO_2 < 85%[¥,**]	229 / 558,459 (0.0)	1,123 / 697,352 (0.2)	0.38 (0.33 to 0.42)[b]	< 0.001

** until extubation and for at least 30 consecutive seconds*

*** until extubation and reported according to the total number of breaths*

[¥] when the measured SpO_2 had a quality index > 50%

[a] effect estimate is mean difference

[b] effect estimate is risk ratio

Table S5. Co-primary and secondary outcomes in the *per*-protocol population

	Fully automated ventilation (*n* = 101)	Conventional ventilation (*n* = 98)	Effect Estimate (95% CI)	*p value*
Primary outcome				
Percentage of time in the optimal zone[**]	57.5 ± 27.0	24.3 ± 29.3	33.2 (25.3 to 41.1)[a]	< 0.001
Median (IQR)	63.8 (37.7 – 80.6)	5.1 (0.0 – 48.0)		
Percentage of time in the acceptable zone[**]	15.4 ± 15.0	51.4 ± 35.2	-36.0 (-43.6 to -28.4)[a]	< 0.001
Median (IQR)	11.9 (4.5 – 21.2)	49.1 (18.4 – 90.3)		
Percentage of time in the critical zone[**]	0.2 ± 0.6	3.0 ± 8.4	-2.8 (-4.5 to -1.1)[a]	0.001
Median (IQR)	0.0 (0.0 – 0.0)	0.3 (0.0 – 1.1)		
Secondary outcomes				
Percentage of breaths in the optimal zone[*]	152,819 / 212,837 (71.8)	61,738 / 197,449 (31.3)	2.32 (2.30 to 2.34)[b]	< 0.001
Percentage of breaths in the acceptable zone[*]	55,882 / 212,837 (26.3)	124,906 / 197,449 (63.3)	0.45 (0.44 to 0.45)[b]	< 0.001
Percentage of breaths in the critical zone[*]	4,136 / 212,837 (1.9)	10,805 / 197,449 (5.5)	0.52 (0.51 to 0.54)[b]	< 0.001
Time until spontaneous breathing, minutes[£££]	165.9 ± 103.9	225.0 ± 175.2	1.49 (1.11 to 2.00)[c]	0.007
Median (IQR)	142 (92 – 219)	174 (121 – 276)		
Duration of weaning, minutes[£]	203.4 ± 204.0	278.5 ± 451.6	1.21 (0.90 to 1.62)[c]	0.203
Median (IQR)	138 (73 – 242)	155 (75 – 325)		
Duration of ventilation, minutes[****]	349.8 ± 240.2	452.7 ± 483.5	1.29 (0.96 to 1.84)[c]	0.084
Median (IQR)	294 (198 – 409)	317 (206 – 548)		
Proportion of failed extubations[***]	2 / 96 (2.1)	2 / 92 (2.2)	0.98 (0.34 to 2.63)[b]	0.999[e]
Incidence of postoperative pulmonary complications	82 / 101 (81.2)	84 / 98 (85.7)	0.86 (0.62 to 1.19)[b]	0.390
Pneumonia	6 / 101 (5.9)	5 / 98 (5.1)	1.08 (0.62 to 1.88)[b]	0.795
Pneumothorax	7 / 101 (6.9)	6 / 98 (6.1)	1.06 (0.63 to 1.80)[b]	0.817
Atelectasis	82 / 101 (81.2)	83 / 98 (84.7)	0.89 (0.63 to 1.24)[b]	0.639
ICU length of stay, days	0.6 ± 0.6	0.7 ± 0.7	-0.1 (-3.4 to 1.6)[a]	0.172
Median (IQR)	0.3 (0.3 – 0.6)	0.4 (0.3 – 0.7)		

Table S5. Continued

	Fully automated ventilation (*n* = 101)	**Conventional ventilation (*n* = 98)**	**Effect Estimate (95% CI)**	***p value***
ICU readmissions[£££]	2 / 101 (2.0)	3 / 98 (3.1)	0.78 (0.26 to 2.31)[b]	0.679[e]
Mortality				
ICU	3 / 101 (3.0)	0 / 98 (0.0)	---[d]	0.246[e]
30-day	3 / 101 (3.0)	0 / 98 (0.0)	---[d]	0.086[f]
Percentage of breaths with SpO_2 < 85%[¥,**]	68 / 216,029 (0.0)	782 / 221,068 (0.4)	0.16 (0.12 to 0.20)[b]	< 0.001

[*] *during the first three hours of ventilation or until extubation and for at least 30 consecutive seconds*

[**] *during the first three hours of ventilation or until extubation and reported according to the total number of breaths*

[***] *defined as any re-intubation within 48 hours after extubation and considering only patients who survived and did not undergo a re-sternotomy during this time*

[****] *time from ICU admission until first successful extubation*

[£] *time from stopping sedatives and a rectal temperature ≥ 35.5 ºC until successful extubation*

[££] *time from ICU admission until ≥ 5 consecutives spontaneous breaths*

[£££] *during the first 72 hours after ICU discharge*

[¥] *when the measured SpO_2 had a quality index > 50%*

[a] *effect estimate is mean difference*

[b] *effect estimate is risk ratio*

[c] *effect estimate is hazard ratio*

[d] *since no event was observed in one group, the effect estimate was not calculated (infinite in the upper limit)*

[e] *calculated with Fisher exact test*

[f] *calculated with log-rank test*

Table S6. Co-primary and secondary outcomes for all recorded breaths and time in the *per*-protocol population

	Fully automated ventilation (*n* = 79)	**Conventional ventilation (*n* = 85)**	**Effect Estimate (95% CI)**	***p value***
Co-primary outcome				
Percentage of time in the optimal zone*	50.9 ± 25.2	25.6 ± 23.9	25.3 (17.8 to 32.9)[a]	< 0.001
Median (IQR)	54.1 (32.4 – 71.3)	18.5 (1.9 – 44.7)		
Percentage of time in the acceptable zone*	15.2 ± 13.3	40.1 ± 29.0	-24.9 (-32.0 to -17.9)[a]	< 0.001
Median (IQR)	11.9 (5.1 – 20.6)	31.4 (15.7 – 66.4)		
Percentage of time in the critical zone*	0.7 ± 3.2	4.0 ± 8.7	-3.2 (-5.3 to -1.2)[a]	0.002
Median (IQR)	0.0 (0.0 – 0.3)	0.6 (0.0 – 2.6)		
Secondary outcomes				
Percentage of breaths in the optimal zone**	267,537 / 381,218 (70.2)	144,480 / 379,043 (38.1)	1.98 (1.97 to 1.99)[b]	< 0.001
Percentage of breaths in the acceptable zone**	100,687 / 381,218 (26.4)	205,668 / 379,043 (54.3)	0.53 (0.52 to 0.54)[b]	< 0.001
Percentage of breaths in the critical zone**	12,994 / 381,218 (3.4)	28,895 / 379,043 (7.6)	0.60 (0.59 to 0.61)[b]	< 0.001
Percentage of breaths with SpO_2 < 85%[¥,**]	157 / 388,504 (0.0)	841 / 498,575 (0.2)	0.35 (0.31 to 0.41)[b]	< 0.001

** until extubation and for at least 30 consecutive seconds*

*** until extubation and reported according to the total number of breaths*

[¥] when the measured SpO_2 had a quality index > 50%

[a] effect estimate is mean difference

[b] effect estimate is risk ratio

Table S7. Additional sensitivity analyses

	Effect Estimate (95% CI)	*p* value
Bootstrapped *t*-test		
For the first three hours of ventilation or until extubation		
Percentage of time in the optimal zone	29.6 (18.6 to 40.5)[a]	< 0.001
Percentage of time in the acceptable zone	-33.1 (-43.3 to -22.8)[a]	< 0.001
Percentage of time in the critical zone	-2.6 (-4.7 to -0.3)[a]	0.025
For all recorded time		
Percentage of time in the optimal zone	21.1 (11.5 to 30.5)[a]	< 0.001
Percentage of time in the acceptable zone	-23.4 (-32.1 to -15.1)[a]	< 0.001
Percentage of time in the critical zone	-2.1 (-4.4 to 0.3)[a]	0.092
Adjusted Cox proportional hazard model*		
For the intention-to-treat population		
Time until spontaneous breathing	1.37 (1.03 to 1.82)[b]	0.028
Duration of weaning	1.32 (0.99 to 1.75)[b]	0.058
Duration of ventilation	1.28 (0.97 to 1.71)[b]	0.080
For the *per*-protocol population		
Time until spontaneous breathing	1.44 (1.07 to 1.94)[b]	0.015
Duration of weaning	1.30 (0.96 to 1.75)[b]	0.087
Duration of ventilation	1.30 (0.96 to 1.75)[b]	0.090

[a] *effect estimate is mean difference*

[b] *effect estimate is hazard ratio*

** adjusted by the duration of surgery, total amount of intravenous sedatives and time from ICU admission until a rectal temperature >35.5°C*

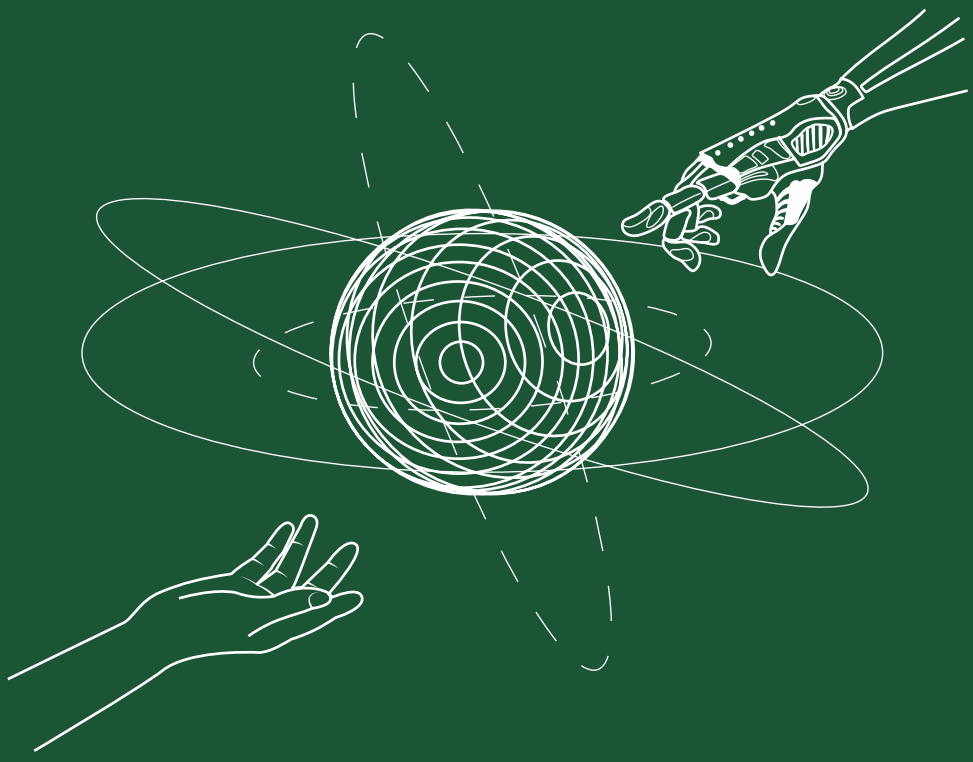
Table S8. Additional post hoc analyses

	Automated (*n* = 109)	Conventional (*n* = 111)	Effect Estimate (95% CI)	*p value*
During the first three hours of ventilation or until extubation				
Percentage of breaths with Pmax ≤ 30 cmH_2O	231,706 / 231,924 (99.9)	249,196 / 251,440 (99.1)	5.44 (4.79 to 6.18)	< 0.01
Percentage of breaths with SpO_2 < 90%*	989 / 232,211 (0.4)	1,758 / 252,244 (0.7)	0.75 (0.71 to 0.79)[a]	< 0.01
Percentage of breaths with SpO_2 > 97%*	77,873 / 232,211 (33.5)	146,079 / 252,244 (57.9)	0.59 (0.58 to 0.59)	< 0.01
Percentage of breaths with SpO_2 ≥ 90% and ≤ 97%¥	139,124 / 232,211 (59.9)	71,362 / 252,244 (28.3)	1.94 (1.93 to 1.96)	< 0.01
For all recorded breaths				
Percentage of breaths with Pmax ≤ 30 cmH_2O	556,477 / 556,995 (99.9)	690,319 / 694,375 (99.4)	3.94 (3.63 to 4.27)	< 0.01
Percentage of breaths with SpO_2 < 90%*	1,469 / 558,459 (0.3)	5,709 / 697,352 (0.8)	0.46 (0.44 to 0.48)	< 0.01
Percentage of breaths with SpO_2 > 97%*	166,218 / 558,459 (29.8)	274,967 / 697,352 (39.4)	0.78 (0.77 to 0.79)	< 0.01
Percentage of breaths with SpO_2 ≥ 90% and ≤ 97%¥	309,324 / 558,459 (55.4)	223,877 / 697,352 (32.1)	1.68 (1.67 to 1.69)	< 0.01
During controlled mechanical ventilation				
Percentage of breaths with Pmax ≤ 30 cmH_2O	179,725 / 179,794 (99.9)	216,451 / 218,630 (99.0)	14.78 (11.71 to 18.64)	< 0.01
Percentage of breaths with Driving pressure ≤ 15 cmH_2O**	171,076 / 172,669 (99.1)	174,991 / 187,358 (93.4)	4.33 (4.13 to 4.53)	< 0.01
Percentage of breaths with Mechanical power < 17 J min^{-1}**	172,546 / 172,643 (99.9)	185,625 / 187,296 (99.1)	8.78 (7.23 to 10.65)	< 0.01
Percentage of breaths with SpO_2 < 90%*	697 / 179,820 (0.4)	1,428 / 219,319 (0.7)	0.72 (0.68 to 0.77)	< 0.01
Percentage of breaths with SpO_2 > 97%*	62,514 / 179,820 (34.8)	133,982 / 219,319 (61.1)	0.55 (0.54 to 0.55)	< 0.01
Percentage of breaths with SpO_2 ≥ 90% and ≤ 97%¥,*	107,825 / 179,820 (60.0)	57,666 / 219,319 (26.3)	2.11 (2.10 to 2.13)	< 0.01
During assisted mechanical ventilation				
Percentage of breaths with Pmax ≤ 30 cmH_2O	51,981 / 52,130 (99.7)	32,745 / 32,810 (99.8)	0.88 (0.81 to 0.96)	0.03
Percentage of breaths with SpO_2 < 90%*	292 / 52,294 (0.6)	330 / 32,909 (1.0)	0.76 (0.70 to 0.83)	< 0.01
Percentage of breaths with SpO_2 > 97%*	15,359 / 52,294 (29.4)	12,096 / 32,909 (36.8)	0.87 (0.86 to 0.88)	< 0.01
Percentage of breaths with SpO_2 ≥ 90% and ≤ 97%¥,**	31,298 / 52,294 (59.9)	13,694 / 32,909 (41.6)	1.33 (1.32 to 1.35)	< 0.01

** when the measured SpO_2 had a quality index > 50%*

*** Driving pressure = estimated Pplateau by the ventilator – PEEP*

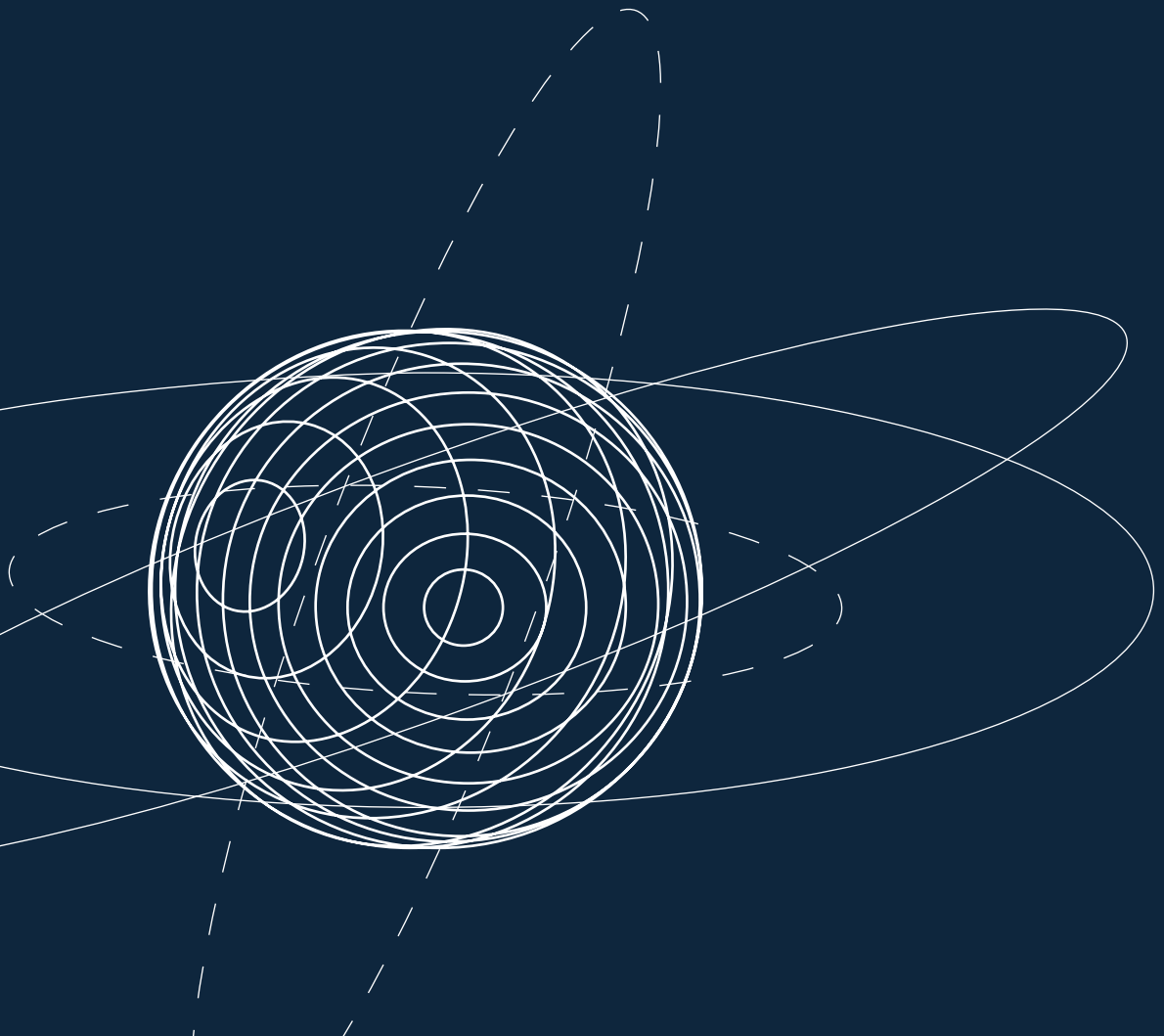
[a] effect estimate is risk ratio



PART 6

Discussion
English summary

6.1



General discussion

Clinical decision support (CDS) systems can assist healthcare professionals in the clinical decision-making process. This support becomes more relevant as more and more patients present with complex combinations of conditions that require at times conflicting therapeutic strategies. To enhance patient care and safety across many clinical domains, numerous clinical decision support (CDS) systems of varying complexity have been developed.[1-4] However, the large-scale integration and acceptance of these systems in clinical workflows is lagging behind due to several challenges that reflect the complexity of clinical practice and the often immature nature of the supporting systems.[4-7] In addition, scientific evidence that defines the requirements of a valuable CDS system is scarce. These requirements are currently based on experience-based recommendations, such as the "Ten commandments of Clinical Decision Support".[2, 4, 7] Clinical evaluation of a CDS system, even early in the development phase, is warranted and key to understanding the advantages and barriers of these CDS systems.[2, 7]

The aim of this thesis was to evaluate the effectiveness of four new computerized CDS systems within the discipline of acute and critical care. This thesis is divided into four parts based on the degree of complexity and level of automation of the investigated CDS systems:

1. A straightforward computerized CDS system without automation
2. A CDS system that is based on a simple model which can gather some data automatically
3. A CDS system based on a complex transparent model with an automated, closed-loop, functionality
4. A fully closed-loop automated CDS system

The key findings of these studies will be outlined to address the primary objective of this thesis, followed by a report of lessons learned, methodological considerations, and future perspectives.

KEY FINDINGS

1. A straightforward computerized CDS system without automation

Can a Crisis Checklist Application on general wards support the teamwork performance of physicians and nurses and acute care management of a deteriorating severely ill patient?

A step-up approach research model was used for the development and clinical evaluation of the Crisis Checklist App. An international survey was conducted to explore which clinical situations clinicians perceive to be a clinical crisis.(**Chapter 2.1**) These insights were used to create the scenarios for a feasibility study (**Chapter 2.2)** and a multicentre simulation-based study (**Chapter 2.3**), both of which evaluated the effects on teamwork performance and acute care management. The key findings of these studies were as follows:

- The opinions of doctors and nurses differ on illness severity and on the need for urgent treatment or escalation of care. In most cases with significantly deranged vital signs (i.e. MEWS ≥7), only a minority of the participating doctors and nurses perceived these cases as a "clinical crisis" (35%; n = 127 of 354) and 66% of the participants decided not to call the rapid response team immediately.(**Chapter 2.1**)
- Both the feasibility study and the multicentre simulation-based study showed improved teamwork performance of the participating teams in the scenarios with the Crisis Checklist App compared to the scenarios without this application. (**Chapter 2.2 & 2.3**)
- The use of the Crisis Checklist App was associated with
- Improvement of non-technical skills in the domains of "situational awareness" and "decision making" based on the assessments of the observers and the participant's self-assessments.(**Chapter 2.3**)
- Reduced percentages of omission of predefined safety-critical steps and reduced time to complete these steps.(**Chapter 2.3**)
- A large proportion of acute and emergency care in hospital care is delivered by junior clinicians in the first years of their training.[8, 9] Acute care management of deteriorating severely ill patients tended to improve especially in teams with less experienced team leaders, e.g. medical students. This finding suggests a key role for computerized checklist aids as a vital educational tool that can prepare and support junior clinicians in dealing with these stressful situations.(**Chapter 2.3**)

2. A CDS system that is based on a simple model which can gather some data automatically

Can an automated MEWS system augment the reliable acquisition of vital signs in patients on general wards in order to improve the protocol-adherence and care management of the bedside clinician?

Part two of this thesis focussed on evaluating the effectiveness of a commercially available automated modified early warning scoring (MEWS) system (Philips Guardian®). The effectiveness of this CDS system was retrospectively analysed in a before-and-after study on a surgical high-dependency unit. The key findings of this study were as follows:

- The implementation of an automated MEWS system in a surgical high-dependency unit improves the recording of vital signs, resulting in more complete MEWS assessments and a better adherence to the local protocol.(**Chapter 3.1**)
- The improved registration of vital signs and protocol adherence tended to reduce 28-day mortality and ICU readmission rates. However, the present study found no statistically significant differences on these and/or any other patient-centred outcomes.(**Chapter 3.1**)

3. A CDS system based on a complex transparent model with an automated, closed-loop, functionality

Can a dynamic clinical checklist of TraceBook improve the adherence of physicians to best eligible practice and their acceptance of checklists during ward rounds in the Intensive Care?

In this part of the thesis, a step-up approach research model was applied as part of the development process of a new CDS system called TraceBook. The aim of this approach was to clinically evaluate the effectiveness, user acceptance, and potential barriers of applying TraceBook's dynamic clinical checklist during its development. The step-up approach started with a simulation-based study followed by a prospective before-and-after mixed method study in real clinical practice on an ICU. The gained knowledge was used to improve the CDS systems usability. The most relevant outcomes of these two studies were as follows:

Quantitative outcomes.

- In both studies the completion rates of paper checklist items in the control groups were similar (± 75%) and comparable to the findings of other studies that evaluated paper checklists.[5-8](**Chapter 4.1 & 4.3**)

- In both studies, this completion rate was improved (90 - 100%) by the implementation of TraceBook's dynamic clinical checklists.(**Chapter 4.1 & 4.3**)
- The availability of TraceBook's dynamic clinical checklist during daily ward rounds on the ICU in the before-and-after study was associated with a reduced ICU length of stay, fewer days with use of empiric antibiotics, and fewer cases with unacceptable levels of critical care pain scores. No reductions were found in mortality rates, daily need of intravenous sedatives, and registration of pre-defined complications. (**Chapter 4.3**)

Qualitative outcomes:

- In the simulation study, the physicians rewarded the dynamic clinical checklist with a high satisfaction score of 4.1 out of 5 points. All physicians considered the dynamic clinical checklist to have the potential for being successfully integrated into daily practice and preventing complications. However, the usability must be optimized to meet these expectations.(**Chapter 4.1 & 4.3**)
- An in-depth qualitative analysis in the mixed-method study showed that participants valued the dynamic clinical checklist as an attractive and innovative tool that could easily be applied in daily practice. Physicians especially appreciated the support that was provided for items related to comprehensive and complicated guidelines. However, they also indicated that TraceBook's checklist still functioned too much as a checklist for care processes instead of as a cognitive aid supporting the more complex decision making at the bedside. (**Chapter 4.3**)

4. A fully closed-loop automated CDS system

Is a fully automated ventilation mode (INTELLiVENT-ASV) able to optimize ventilatory support in patients admitted to Intensive Care?

In this part, we first evaluated the current clinical use of closed-loop artificial ventilation modes on intensive care units in the Netherlands.

- Of the 72 interviewed representatives of Dutch non-paediatric ICUs, 58% reported to have access to a closed-loop ventilation mode. In 43% of these ICUs, a closed-loop ventilation mode was applied frequently (i.e. at least once in the week the interview).(**Chapter 5.1**)
- Reasons for not using the closed-loop ventilation modes were lack of knowledge (40%), insufficient evidence reporting a beneficial effect (35%) and lack of confidence (25%). Of all the respondents, 10% spontaneously added that a perceived lack of control might also play a role for not using these modes.(**Chapter 5.1**)

Secondly, a non-inferiority prospective study (**Chapter 5.2**) evaluated the safety and effectiveness of the most advanced automated ventilation mode currently available in clinical practice, the INTELLiVENT-ASV. This automated ventilation mode was compared with two other ventilation strategies in postoperative cardiac surgery patients following a fast-track pathway: 1) A conventional ventilation strategy with controlled mandatory ventilation (CMV) followed by pressure support ventilation (PSV); and 2) Assisted Support Ventilation (ASV), the predecessor of INTELLiVENT-ASV.

The most important outcomes of comparing these three ventilation strategies were as follows:

- No ventilation-related safety issues requiring interventions were observed. (**Chapter 5.2**)
- The use of INTELLiVENT-ASV was associated with a smaller number of registered interactions with the ventilator than the use of either of the two other strategies. The number of registered interactions with the ventilator when applying ASV was also reduced compared to the conventional strategy.(**Chapter 5.2**)
- The use of INTELLiVENT-ASV did not translate into a shorter median postoperative mechanical ventilation time compared to either of the two other strategies. (**Chapter 5.2**)

Finally, a randomized controlled trial (**Chapter 5.3**) was conducted to compare automated ventilation, INTELLiVENT-ASV, with a lung protective conventional ventilation strategy (CMV followed by PSV) in non-fast-track patients after cardiac surgery. The findings of this study can be summarized as follows:

- INTELLiVENT-ASV achieved its intended purpose to reduce the time of ventilatory support in a predefined undesired critical zone of ventilation compared to the control group.(**Chapter 5.3**)
- The fully automated ventilation mode favourably changed time spent in the optimal and acceptable zones of ventilation, reduced the number of breaths with severe hypoxemia (SpO_2 < 85%), and shortened the time until spontaneous breathing. (**Chapter 5.3**)
- Secondary outcome analyses showed no beneficial effect of automated ventilation on other patient-centred outcomes for which this study was not powered. In other words, there was no significant difference between the automated and conventional ventilation strategy in terms of the total duration of ventilation and weaning, the proportion of patients with failed extubation or postoperative pulmonary complications, ICU length of stay or mortality.(**Chapter 5.3**)

- The favourable outcomes were found in a group of patients that was deliberately chosen for their limited number of residual confounders, such as the absence of major lung injury. This study design implies that the favourable outcomes are attributable to the intervention, even in the relatively short period of recorded post-operative mechanical ventilation and in a centre with well–trained and experienced ICU nurses and doctors. Larger differences might therefore be expected in patients with more severe lung diseases, in resource–poor training facilities or in other settings with fewer resources or fewer staff members.(**Chapter 5.3**)

LESSONS LEARNED

The key findings of the research reported in this thesis showed that all four CDS systems fulfilled their intended clinical purposes. This research has provided insights into the conceptual development and evaluation of CDS systems in general practice and specifically in acute and critical care. These insights can assist future developers, researchers and healthcare professionals in their efforts to develop, study, implement, and choose a useful CDS system for their clinical practice.

CDS systems in acute and critical care: development and evaluation

From development to market release

The development of a CDS systems until its market release into clinical practice is a complex process which requires collaboration and bridging gaps between engineers, scientists, and the end-users (*Figure 1*). Based on the experience gained during the creation of The Crisis Checklist App (**Part 2**) and TraceBook (**Part 4**), interprofessional teams, sometimes including patients, are encouraged to start collaboration early so that all collaborators understand the CDS system's clinical purpose and core functionality, the technical limitations, and the context of use with corresponding pitfalls. These teams also need to consider that a health-related decision making process involves at least two persons: the patient and the healthcare professionals.[2] This multi-person involvement further complicates the process of developing and evaluating new CDS systems because the interests of all parties need to be protected. These interests include the patient's quality of life and survival (patient-centred outcomes) and the healthcare professionals' quality of care, workflow, and working environment (user-centred outcomes), which can all be negatively influenced by various technological, design-related, clinical, and socio-organizational barriers.[2-4, 7]

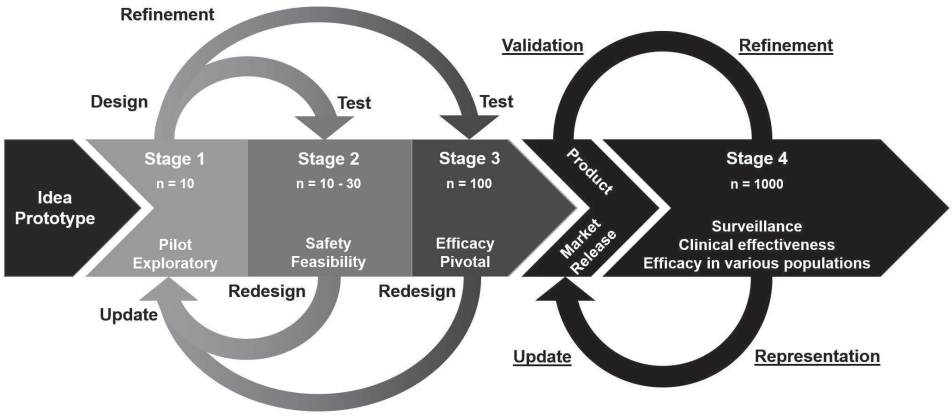


Figure 1. From idea to post market evaluation with the clinical trial stage sequence for evaluating new CDS systems

A difference between the development and implementation of a CDS system and the development and introduction of a new drug is that the system's functionalities and user-interface can be tweaked and improved, even after the market release, until the system matches its clinical purpose and the requirements of the intended end-users. Therefore, outcome measures of the clinical evaluations of CDS systems must represent the interests of all parties and should serve to refine the CDS system (*Figure 1*). This was the case during the development of the Crisis Checklist App and TraceBook's dynamic checklist.(**Part 2 & 4**)

The evaluation of a CDS system

The method of clinically evaluating CDS systems is fairly similar to the clinical trial phase sequence for testing new drugs (*Figure 1*). Similar to the evaluation of a drug, the primary aim of the early trials is demonstrating the safety of a CDS system and of its ability to produce an intended clinical purpose (effectiveness), e.g. enhancing the reliable acquisition of medical data or increasing checklist completion rates.(**Chapter 3.1, 4.1 & 4.3**) After all, a CDS system must achieve a safe improvement of the current clinical practice to show a favourable effect on patient- or user-centred outcomes. In these trials, outcomes should be measured in the intended context of use, or in a close representation of this intended context, with residual confounders being reduced as much as possible to demonstrate that the findings are attributable to the CDS system. If in these settings the CDS system is shown to be safe and effective, larger clinical trials can be organized to assess the CDS system's short- and long-term effects on patient-centred and/or user-centred outcomes (*Figure 1*).

In medical research, the highest level of evidence for evaluating these outcomes would be generated by means of randomized controlled trials.[10, 11] However, the use of this study design for evaluating CDS systems in clinical settings is hampered by the nature of the intervention, which precludes the randomization of patients, the blinding of healthcare professionals, and the use of placebos.[12] In acute and critical care, the use of this study design is hindered even more by the presence of multiple confounders in critically ill patients and by the unpredictable character of clinical emergencies. These challenges highlight the importance of carefully planning a well-controlled study design. The experiences gained in preclinical trials and in the earliest smaller clinical trials can help to plan such well-controlled studies.

For this thesis, three feasibility studies were performed, two of which were simulation-based in a preclinical stage.(**Chapter 2.2 & 4.1**) Simulation was perceived to be an excellent choice for preclinical testing because this strategy enabled the developers and researchers to observe the end-users using the CDS system in safe simulated environments that represent the context of use.(**Chapter 2.2 & 4.1**). The findings of these studies facilitated further improvements of the CDS systems; their designs were deemed feasible, which helped to design a larger multinational simulation study and a before-and-after study in real practice.(**Chapter 2.3 & 4.3**)

CDS SYSTEMS IN ACUTE AND CRITICAL CARE: IMPORTANT FUNCTIONALITIES

Healthcare professionals working in acute and critical care need to make numerous decisions, sometimes rapidly, for severely ill patients in intense and distracting environments with various medical devices providing a large amount of information.[13-15] The research of this thesis generated a number of interesting key lessons related to CDS system functionalities, which complement the existing set of recommendations (*Figure 2*).

Improving the human-technology interaction: new digital input

Digital medical data repositories are still largely depending on the data input from healthcare professionals which makes these datasets vulnerable if this input is incomplete or even absent. Accordingly, CDS systems depending on these data repositories cannot ascertain if and why a certain input is absent. These CDS systems can only verify that a certain issue was considered by the healthcare professionals if data changes occur in the repository. For instance, an elevated pain score raises the issue that the treatment must be changed based on local protocol. If nothing changes in the data

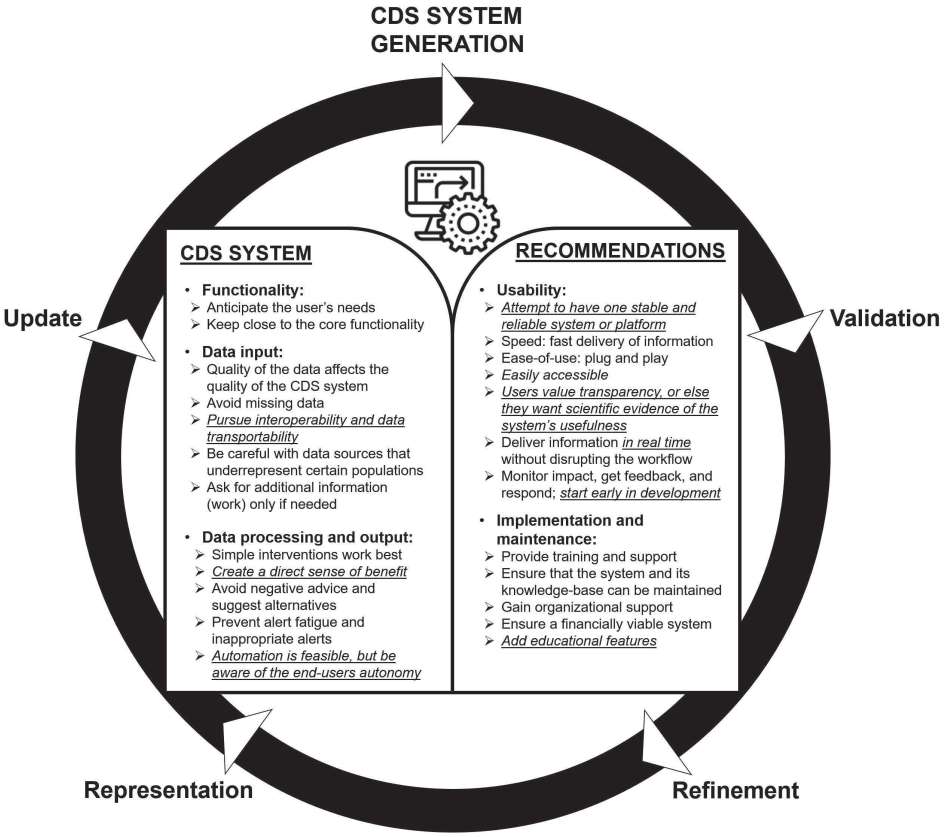


Figure 2. Complemented set of the experience-based recommendations for a useful CDS system in acute and critical care

A complemented set of experience-based recommendations for a useful CDS system based on various technological, design-related, clinical, and socio-organizational challenges that reflect the complexity of clinical practice. The life cycle surrounding the CDS system recommendation represents the continuous need to update CDS systems.[2-7, 20] *The underlined text in italics was added based on the findings of the research described in this thesis.*

repository, the system cannot ascertain whether the healthcare professional forgot to consider this issue, or whether the healthcare professional deliberately chose not to intervene. TraceBook's dynamic checklists, on the other hand, can notice missing data and confront the caregiver directly with these issues. The dynamic checklist's content shows which care processes deviate from the locally agreed protocolled care. The user can then decide to change to the agreed practice, which is then digitally recorded (e.g. changing the dosage of a drug or registering a new observation), or he or she can check

that this issue was considered and even provide a reason for the deviation. Adding this information will also provide important insights for retrospective data analyses and can help to administer and track why certain clinical decisions were made, especially in more complex clinical situations. This aspect may prove to be an important aspect of a CDS system because it enables the system to analyse and learn if and when deviations from protocols are beneficial for patients in specific complex situations.

Effective communication between systems

In critically ill patients, multiple organ systems are affected that dynamically interact with each other, e.g. the heart and lungs. In acute care and critical care, these organ systems and their interactions are periodically or continuously monitored by means of multiple medical devices, diagnostic imaging, laboratory results, and upcoming tools such as biosensors. The use of all these technologies exposes healthcare professionals to growing amounts of medical data, including enormous amounts of raw data and wave patterns that cannot be processed nor interpreted by humans.[3, 14] Advanced CDS systems are now being developed that are able to combine, structure and leverage this data, such as non-knowledge based CDS systems that use artificial intelligence, machine learning, or statistical pattern recognition. However, the creation of multiple stand-alone CDS systems with various user-interfaces and user-experiences is undesirable for healthcare professionals that already have to deal with complex, time-pressured and distracting conditions. The value of upcoming CDS systems and medical devices can therefore be enhanced if these systems are interoperable, or able to communicate and transport data effectively to other devices. For example, the capacity of TraceBook to gather medical data from various data sources made it possible to support the decision-making process concerning multiple care processes, instead of having multiple stand-alone CDS systems.(**Part 4**) On the other hand, the effectiveness of these systems is very likely to have been influenced negatively by the administrative burden of copying all the results from the automated MEWS system to the Electronic Medical Records. (**Chapter 3.1**)

Respect autonomy and transparency

The ICU is a highly controlled environment, where healthcare professionals attempt to control and monitor all parameters because they are responsible for the medical decisions concerning their vulnerable patients. This responsibility and the healthcare professionals' desire of autonomy makes it difficult to entrust the decision-making process to a machine, especially if this machine has a complex and incomprehensible architecture or algorithm. Healthcare professionals in acute and critical care therefore highly value a CDS system with a comprehensible development process and a

transparent model of the decision-making process.[2, 16] In this thesis, the importance that healthcare professionals attribute to their autonomy and to a system's transparency also emerged from the questionnaires and from the interviews conducted in the studies that evaluated the usability and acceptance of the more complex CDS systems with closed-loop functionalities. Healthcare professionals reasoned that perceptions of "lack of control" and "lack of trust" were important barriers that hamper the use and implementation of closed-loop ventilation systems.(**Chapter 5.1**) Others argued that clinicians should be cautious of becoming too dependent on these technologies, as this dependence might impair the clinicians' knowledge and clinical reasoning skills. (**Chapter 4.3**) The emergence of more advanced CDS systems with non-knowledge-based models and closed-loop functionalities will reinforce these sentiments if the effects on autonomy and the CDS system's transparency are not being duly considered during the processes of development and clinical evaluation. These challenges, however, should not stop the development of advanced CDS systems with automated functionalities. Healthcare professionals and patients can certainly benefit from these CDS systems because they are able to reduce the variability of care, to help tailoring evidence-based treatments to patients' needs, and to improve the work environment by reducing the workload of healthcare professionals.(**Part 4 and 5**)

Direct sense of benefit

The acceptance and perceived usefulness of a CDS system are enhanced if its assistance creates a direct sense of benefit. For example, users appreciated TraceBook's functionalities that reminded users of unresolved care processes and that supported their decision-making process by highlighting relevant data from the Electronic Medical Record while providing evidence-based advice from the local guidelines. (**Chapter 4.1 and 4.3**) Young and less experienced physicians and nurses valued the static syndrome-based "read-and-do" checklist of the Crisis Checklist App as it improved their self-confidence and helped them to speak up as a member of the team.(**Chapter 2.3**) By contrast, more experienced healthcare professionals sometimes criticized the content of these checklists or the ICU ward round checklists of TraceBook because in their opinion several steps or items were too straightforward and were already taken care of without the checklists. This observation implies that a dynamic functionality of a CDS system must not only tailor the content to the care of each specific individual patient, but must also learn to tailor this content to the healthcare professional's level of experience. More research is needed to explore if this dynamic functionality can further improve the acceptance of and compliance with checklists or CDS systems.

The direct sense of benefit can also be accomplished by providing an educational carry-over effect. For instance, less experienced clinicians indicated that the automated EWS system helped them to improve their clinical reasoning skills for deteriorating patients, (**Chapter 3.1**) that the syndrome-based checklist from the Crisis Checklist App taught them which acute care management steps are essential,(**Chapter 2.2 and 2.3**) and that TraceBook's ability to display the relevant local guidelines and references helped them to obtain more in-depth knowledge.(**Chapter 4.1 and 4.3**) The use of gamification was a more unexpected element of a CDS system that could make the use of the CDS system more satisfactory. For example, TraceBook used progress bars to track if all items were completed or if new unresolved items occurred. Users indicated that this functionality added a game-element to the CDS system which encouraged their use of the system.

METHODOLOGICAL CONSIDERATIONS

The studies within this thesis are based on several types of research study designs, each with its inherent strengths and limitations. The limitations of these studies need to be taken into consideration when interpreting their results. Some important generic limitations that were applicable to multiple studies will be addressed here for some extra attention.

Firstly, safety and effectiveness outcome measurements related to the intended purpose of the CDS system were chosen as primary criteria to evaluate the four CDS systems. No firm conclusions can therefore be drawn with respect to benefits regarding patient-centred outcomes. Even though favourable patient-centred outcomes were observed in the studies performed in real practice,(**Chapters 3.1, 4.3, 5.2 & 5.3**) none of these studies were designed or powered to investigate these outcomes.

Secondly, caution is needed when extrapolating the results to other patient categories than the populations that were included in the studies. For example, to evaluate the effectiveness of an automated ventilation mode, we deliberately chose patients after cardiac surgery with minor to no lung dysfunction.(**Chapters 5.2 & 5.3**) Although this strategy reduced the influence of residual confounding, which implies that the favourable outcomes are attributable to the intervention, more research is still required to validate these effects and to determine the effects in other patient categories.

Thirdly, it is difficult to compare the results presented in this thesis with the results of published literature since scientific research that addresses these or similar CDS systems

in the chosen contexts is scarce or even missing. This was especially the case for the Crisis Checklist App and TraceBook's dynamic clinical checklist, since the studies in this thesis were the first ones to evaluate these new CDS systems.

Besides, validity evidence of some tools that were used to assess non-technical skills or usability was not available for the acute and critical care contexts in which they were applied.

Finally, some results originated from simulation-based studies. Simulation has a long history in healthcare education and is valued for its ability to reproduce clinical conditions in a safe environment without endangering patients and clinicians.[17-19] Simulation was considered an excellent alternative to test new CDS systems during their development, but it remains a simulation of clinical decision-making under real-world conditions. The favourable effects found in these studies should therefore be seen as hypothesis-generating observations which require research to determine the validity of these finding in real-world clinical settings.

FUTURE PERSPECTIVES

The findings presented in thesis are promising and provide several points of interests for future research.

Reproducibility and external validity

Multiple studies in this thesis were single-centre studies or the first ones to evaluate a new CDS system. Future research is therefore needed to determine the reproducibility of the studies and the generalizability of their findings.

Personalized medicine in CDS systems

Several functionalities of the tested CDS systems can be enhanced to improve their clinical relevance including data transportability, system integrations, and transparency. Paying extra attention to these functionalities is also important for the development of future, more advanced, CDS systems in critical care. All CDS systems examined in this thesis still had a guideline-based architecture, making them susceptible to the limitations of clinical guidelines in critical care. Guidelines are strong tools for standardization, but their applicability in critical care is limited due to the difficulties of attending to the individual needs of critically ill patients, and to the nonlinear nature of many pathologies and interventions related to critical care. These limitations can be overcome by personalization of the CDS system's output. However, the implementation

of personalized medicine will require the aggregation of large datasets containing all kinds of data, varying from continuously generated raw data of vital sign monitoring to structured and even non-structured data (e.g. free-form text). The size of these data repositories is likely to be too large to be collected and interpreted by humans. Advanced CDS systems using machine learning algorithms are therefore needed to obtain and analyse these large datasets in order to provide personalized decision support relevant to the care of each individual patient without disrupting the clinical workflow.

Multidisciplinary collaborations

The chance of achieving the objectives mentioned above can be enhanced by multidisciplinary collaborations of medical professionals, and technical engineers from universities and the industry. Early collaboration in the process of development allows knowledge and experience to be shared while the purposes and expectations are aligned. In addition, this also allows CDS systems to be clinically evaluated early in their development process, which can directly lead to improvements that increase the CDS system's acceptance and effectiveness. The work presented in this thesis was only feasible due to such a collaboration (IMPULS II) between the Catharina Hospital Eindhoven, Eindhoven University of Technology, and Philips Research. However, caution is needed to interpret this collaboration as sufficient evidence for a successful collaboration strategy. Although the benefits seem obvious, this collaboration can only be seen as one successful example. Future evaluation of similar strategies is needed to understand if this strategy is effective for developing and investigating CDS systems.

Future perspectives for each tested CDS system

The clinical evaluations of each CDS system highlighted several opportunities for improvements and future research relevant to each CDS system:

The Crisis Checklist App

- Its worldwide availability and applicability are important advantages of this application, but integration with local systems can help users in stressful situations by highlighting relevant data from the medical databases, such as laboratory results, for extra attention.
- The ease of use can be improved by creating a multi-language application.
- Research of the application in actual clinical practice is still warranted and should focus on patient-centred outcomes, usability, and the cost-benefit ratio.
- Future simulation-based studies should define the role of this application as a vital educative tool.

Automated MEWS systems

- More comfortable wearable devices for monitoring vital signs are now available for continuous monitoring on general wards. Interoperability between these devices with other medical devices, such as the Electronic Medical Record, can enlarge the available dataset for more advanced algorithms in order to improve the reliability of their predictions of clinical deterioration.
- Future research should explore how the obtained data and trends are to be translated to the practice of bedside clinicians. Three potential solutions come to mind:
- The awareness on the wards can be augmented by means of large screens that outline patients in the order of their risk of deterioration.
- Notifications can be pushed to the clinical team responsible for the patient when the trend starts to deteriorate.
- Specialized acute care teams can automatically be consulted when a certain threshold, relevant to the patient, is reached.
- The automated MEWS systems are able to improve clinical reasoning of healthcare workers by visualizing the measured vital signs combined with grading a patient's severity of illness. This educational effect might be a focus of future studies.

TraceBook's dynamic clinical checklist

- Exploration of TraceBook's ability to improve human-technology interaction. The dynamic content of TraceBook's checklist enables the system to request healthcare workers, or even the patient, to register missing data, thus improving the CDS system's output. This process turns an effort into a direct benefit for the user.
- Employing sophisticated algorithms based on machine learning might improve the relevance of the checklist and augment the support for more complex clinical decisions. However, the impact of the use of these algorithms on the system's transparency must be investigated.
- More research is needed to understand the effect of automating clinical decision-making on the user experience and on patient-centred outcomes.
- Future research should investigate the implementation and validation of multiple interacting dynamic clinical checklists for various clinicians within a patient's clinical pathway.
- The educational effect of TraceBook at the point of care must be studied as it requests attention to care processes that deviate from the local guideline while enabling users to review the corresponding guideline recommendation upon request at the point of care.

Fully Automated ventilation modes

- The safety and effectiveness of personalized mechanical ventilation can be optimized by ensuring interoperability with other medical devices, such as monitoring systems, medical imaging databases and Electronic Medical Records.
- A consolidated and quantitative review is warranted to establish the effect of INTELLiVENT-ASV, the most advanced fully closed-loop ventilation mode, on patient-centred outcomes in various groups of patients.
- The effects of automating the process of mechanical ventilation on the situational awareness and education of the bedside clinicians require more research.
- Future studies are needed to determine the cost–effectiveness of fully automated ventilation, both in resource–rich and resource–poor settings

CONCLUSION

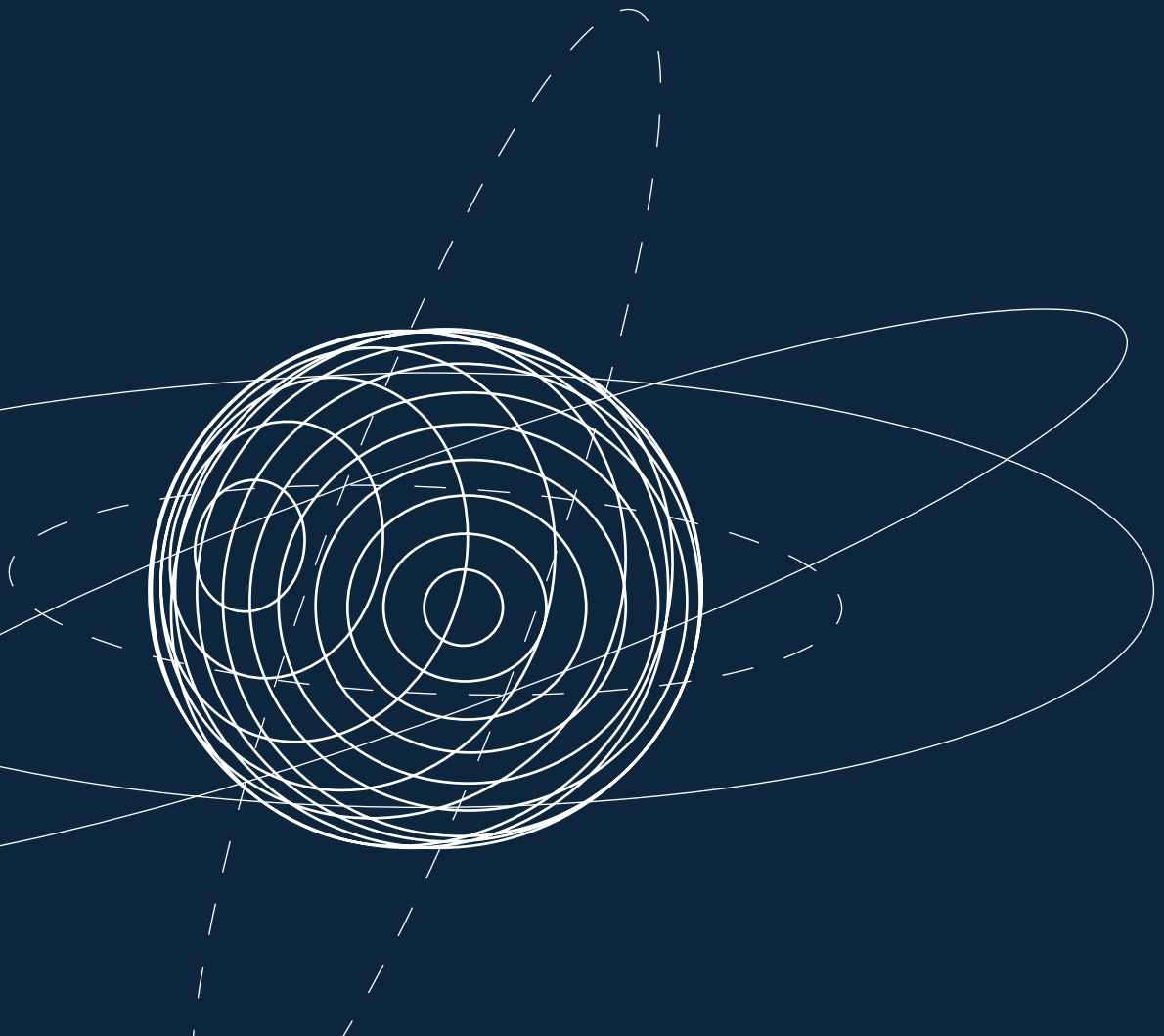
Healthcare professionals in acute and critical care make multiple complex decisions in a timely manner to take care of their severely ill patients. The decision-making process occurs in intense and distracting environments where these clinicians are exposed to large sets of medical data from various medical devices and the Electronic Medical Record. This decision-making process can be supported by CDS systems, but these systems are required to be safe and must be effective. The four investigated CDS systems in this thesis were all safe and all fulfilled their clinical purposes, from improving teamwork and acute care management, generating a larger and more reliable dataset, and increasing compliance with best eligible care to preventing unsafe mechanical ventilation. Although the studies were not designed nor powered to investigate patient-centred outcomes, some CDS systems showed tendencies of improvement in several patient-centred outcomes. This research also provided insights into the conceptual development and evaluation of CDS systems, which contributed to an updated set of recommendations for useful CDS systems in acute and critical care. These insights can assist future developers, researchers and healthcare professionals in their efforts to improve the four tested CDS systems, incorporate some of the investigated functionalities in existing medical devices, or make or choose a new CDS system for clinical practice.

REFERENCES

1. Wasylewicz ATM, Scheepers-Hoeks AMJW. Clinical Decision Support Systems. In: Kubben P, Dumontier M, Dekker A, eds. *Fundamentals of Clinical Data Science*. Cham: Springer International Publishing, 2019; 153-69
2. Bezemer T, de Groot MC, Blasse E, et al. A Human(e) Factor in Clinical Decision Support Systems. *Journal of medical Internet research* 2019; **21**: e11732
3. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digit Med* 2020; **3**: 17
4. Middleton B, Sittig DF, Wright A. Clinical Decision Support: a 25 Year Retrospective and a 25 Year Vision. *Yearbook of medical informatics* 2016; **Suppl 1**: S103-16
5. Shortliffe EH, Sepulveda MJ. Clinical Decision Support in the Era of Artificial Intelligence. *Jama* 2018; **320**: 2199-200
6. Osheroff J, Teich, J. , Levick, D. , Saldana, L. , Velasco, F. , Sittig, D. , Rogers, K. , & Jenders, R. Improving Outcomes with Clinical Decision Support: An Implementer's Guide. 2012
7. Bates DW, Kuperman GJ, Wang S, et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *Journal of the American Medical Informatics Association : JAMIA* 2003; **10**: 523-30
8. Massey D, Chaboyer W, Anderson V. What factors influence ward nurses' recognition of and response to patient deterioration? An integrative review of the literature. *Nursing open* 2017; **4**: 6-23
9. Ludikhuize J, Dongelmans DA, Smorenburg SM, Gans-Langelaar M, de Jonge E, de Rooij SE. How nurses and physicians judge their own quality of care for deteriorating patients on medical wards: self-assessment of quality of care is suboptimal*. *Critical care medicine* 2012; **40**: 2982-6
10. Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. *Jama* 1995; **274**: 1800-4
11. Greenhalgh T. How to read a paper. Getting your bearings (deciding what the paper is about). *BMJ (Clinical research ed)* 1997; **315**: 243-6
12. Armenia S, Thangamathesvaran L, Caine AD, King N, Kunac A, Merchant AM. The Role of High-Fidelity Team-Based Simulation in Acute Care Settings: A Systematic Review. *Surgery journal (New York, NY)* 2018; **4**: e136-e51
13. Lighthall GK, Vazquez-Guillamet C. Understanding Decision Making in Critical Care. *Clinical medicine & research* 2015; **13**: 156-68
14. Ghassemi M, Celi LA, Stone DJ. State of the art review: the data revolution in critical care. *Critical care* 2015; **19**: 118
15. Needham DM. Patient safety, quality of care, and knowledge translation in the intensive care unit. *Respiratory care* 2010; **55**: 922-8

16. Kappen TH, van Loon K, Kappen MA, et al. Barriers and facilitators perceived by physicians when using prediction models in practice. *Journal of clinical epidemiology* 2016; **70**: 136-45
17. Kothari LG, Shah K, Barach P. Simulation based medical education in graduate medical education training and assessment programs. *Progress in Pediatric Cardiology* 2017; **44**: 33-42
18. Cook DA, Hatala R, Brydges R, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *Jama* 2011; **306**: 978-88
19. Lamé G, Dixon-Woods M. Using clinical simulation to study how to improve quality and safety in healthcare. *BMJ Simulation and Technology Enhanced Learning* 2018: bmjstel-2018-000370
20. Greenes RA. Clinical Decision Support. In: Greenes RA, ed. *Clinical Decision Support*. Burlington: Academic Press, 2007; 3-29

6.2



English summary

ENGLISH SUMMARY

Implementing rapidly advancing medical knowledge and technologies into clinical practice is extremely challenging. This is especially true in acute and critical care, which demands quick appropriate decisions to be made on the care of fragile patients in intense and distracting circumstances. Getting a clear overview of the patient's condition, needs and personal desires is hampered by the intrinsic complexity of critical illness and other factors such as the daily variation in personnel, staff experience, work intensity, and medical data overload. Clinical decision support (CDS) systems have been proposed as technologies that should be able to improve complex decision-making and help clinicians deliver better value- and evidence-based personalized care.

Although numerous CDS systems of varying complexity have been developed, their large-scale integration and acceptance into clinical practice has been slow. Apart from the complexity of clinical practice there are several reasons for this, which include the lack of uniform consensus definitions of CDS systems and their requirements (i.e. not all CDS are the same), and sparse scientific evidence of their benefits. The principle type of CDS systems can be based on the degree of complexity and level of automation: static *aid de memoires* such as checklists or algorithms, semi-automatic systems that manipulate manually entered data to generate prompts, reminders and recommendations, and fully automatic "closed-loop" systems that automatically collect data and then manipulate it to regulate a therapeutic intervention. A static non-automatic CDS system that does not manipulate data, even if computerized, is not considered to be a medical device, whereas semi-automatic and automatic CDS systems are. This thesis evaluated the effectiveness of four new computerized CDS systems within the disciplines of acute and critical care. It is divided into four parts based on the degree of complexity and level of automation of the CDS system investigated.

Part 1 describes a static computerized CDS system designed to help first responding clinicians with their acute decision-making and teamwork performance at the bedside of a deteriorating severely ill patient. In these complex, uncertain, and stressful moments clinicians must rely on their pre-existing general knowledge and may fail to notice important details. This is especially so for junior clinicians in the first years of their training, who often provide a large part of many hospitals' acute care. The international survey in **Chapter 2.1** demonstrated that the opinions of doctors and nurses disagree on illness severity and on the need for urgent treatment or escalation of care, which may hamper the success of clinical escalation protocols. **Chapter 2.2** describes a computerized application containing static syndrome-based checklists (Crisis Checklist App). This app was designed to help improve the recognition and acute management

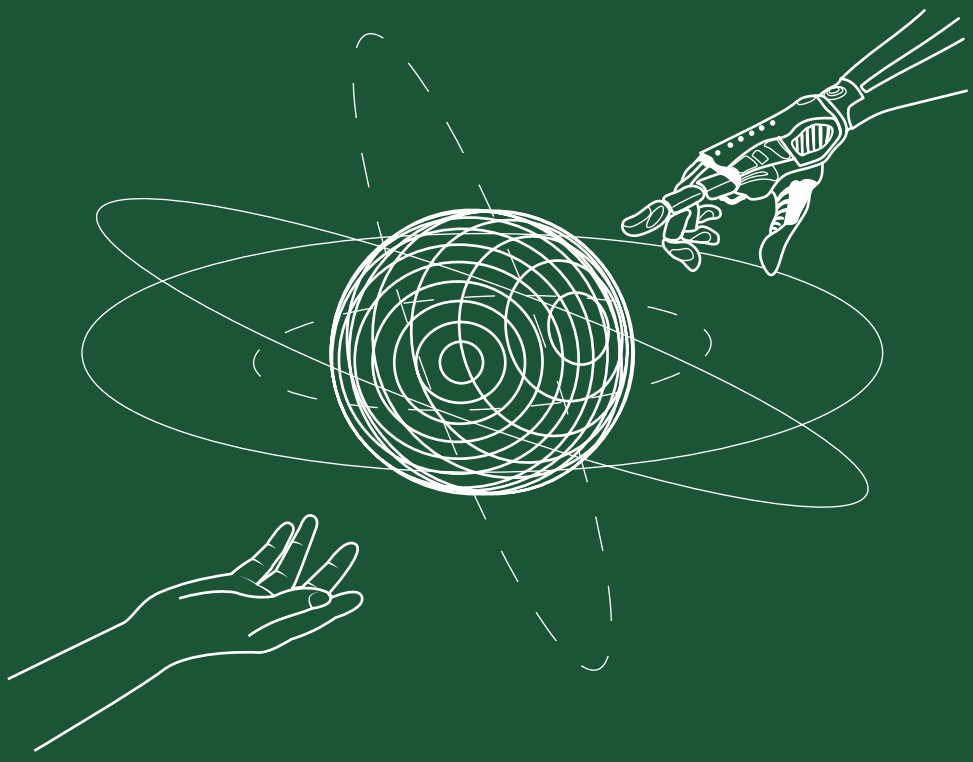
of severely ill patient who were deteriorating. A feasibility study reported in **Chapter 2.2** and a multicentre simulation-based study reported in **Chapter 2.3** demonstrate the potential for this approach to improve acute care management, and non-technical skills such as teamwork performance.

Part 2 describes an electronic CDS system designed to improve the recording of vital signs by automatically and accurately calculating early warning scores, such as the modified early warning score (MEWS). Early warning scores support the early bedside recognition of clinical deterioration and can be used to trigger prompt appropriate early targeted management. However, in clinical practice their effectiveness is limited by incomplete, inaccurate recordings of vital signs, and calculation errors. **Chapter 3.1** reports a retrospective before-and-after implementation study of this automated MEWS system in a surgical high-dependency unit. It showed an increase in the number of complete sets of vital signs recorded, which resulted in more MEWS being calculated and better adherence to local treatment protocols. Although statistical significance was not reached, these improvements reduced 28-day mortality and ICU readmission rates.

Part 3 focusses on the challenge of translating and applying available medical knowledge and evidence into clinical practice. The difficulty in acute and critical care is to tailor the advice of checklists and protocols to the needs of complex critically ill patients with multisystem disease. Often these combinations of conditions may require carefully balancing conflicting therapeutic recommendations and strategies. These patients require a more personalized approach. TraceBook is a complex electronic CDS system, which creates personalized digital checklists containing patient specific recommendations for multiple care processes. **Chapter 4.1** is the first ever report of TraceBook's dynamic clinical checklists in a clinical simulation study. **Chapter 4.2** uses the results of this study to show the importance and added value of TraceBook's functionalities over static and paper-based checklists. **Chapter 4.3** reports a prospective before-and-after mixed method study that evaluated in "real life" the effectiveness and user acceptance of TraceBook's dynamic checklists during ICU ward rounds. The availability of the dynamic clinical checklist during these rounds improved checklist compliance from 75% using a pen and paper-based system to over 90%. This improvement was maintained, even after 8 weeks of use. The availability of these dynamic clinical checklists during ICU ward rounds was associated with a reduced ICU length of stay, shorter courses of empiric antibiotics, and fewer episodes of patients being in unacceptable pain. Most physicians valued the dynamic clinical checklists as an attractive and innovative tool that could easily be applied in daily practice with the potential for preventing complications, and especially appreciated the help provided to implement comprehensive and complicated guidelines (**Chapter 4.1 & 4.3**).

Part 4 examines on a closed-loop mechanical ventilation mode CDS system which automatically tailors the safest ventilator settings in each breath to the patient's needs. Unsafe ventilator settings affect outcomes of critically ill patients with or without pre-existing lung disease. Incorrect ventilator settings over relatively short periods, such as during cardiac surgery, can cause lung injury. **Chapter 5.1** reports a national survey that found that although more than half of the 72 Dutch non-paediatric ICUs have access to a closed-loop ventilation mode, most units (57%) only used the system occasionally or never for the following reasons: lack of knowledge, a perceived lack of confidence and control, and insufficient evidence reporting a beneficial effect. **Chapter 5.2** reports a prospective non-inferiority observational trial in fast-track patients after cardiac surgery and **Chapter 5.3** a randomized controlled trial in non-fast-track patients after cardiac surgery. Both studies showed that the use of a fully closed-loop automated ventilation mode, INTELLiVENT-ASV, was as safe as conventional ventilation modes controlled by well–trained and experienced ICU nurses and doctors. These studies demonstrated that using fully automated ventilation made the provision of lung-protective ventilation more likely, with fewer episodes of severe hypoxaemia, fewer interactions with the ventilator, and a more rapid return to spontaneous breathing.

The four investigated CDS systems in this thesis were all safe and all fulfilled their intended clinical objectives (i.e. improved teamwork performance and acute care management, generated a larger and more reliable dataset of vital signs, optimized compliance with consensus guidelines in complex clinical situations, and ensured lung-protective ventilation). Although the studies were not designed nor powered to investigate patient-centred outcomes, they suggest that CDS systems probably improve several patient-centred and user-centred outcomes. Besides emphasizing the importance of early and close collaboration within an interprofessional team, this research provides insights into the conceptual development and evaluation of CDS systems, which have contributed to an updated set of recommendations for useful CDS systems in acute and critical care.



PART 7: APPENDICES

Publiekssamenvatting
Dankwoord
Curriculum Vitae
Publicatielijst

PUBLIEKSSAMENVATTING

In ons dagelijks leven maken we al veel gebruik van algoritmes die ons helpen met het maken van beslissingen, zoals navigatie-apps en apps met muziekaanbevelingen. Beslissingsondersteunende systemen zijn ook in opkomst in de gezondheidszorg. Deze systemen kunnen vooral zorgverleners (onder meer verpleegkundigen en artsen) in de acute en intensieve zorg ondersteunen omdat zij in aanraking komen met ernstig zieke patiënten. Deze kwetsbare patiënten hebben weinig reserves waardoor het nemen van de juiste beslissingen op het juiste moment cruciaal is. Voor dit proefschrift zijn 4 innovatieve, digitale beslissingsondersteunende systemen voor de acute en intensieve gezondheidszorg onderzocht op hun veiligheid en bruikbaarheid bij de zorg voor ernstig zieke patiënten.

In **hoofdstuk 2** staat de herkenning en eerst behandeling van ernstig zieke patiënten door zorgverleners op reguliere ziekenhuisafdelingen centraal. De mening van deze zorgverleners over de ernst van de ziekte en de benodigde behandeling bleken sterk te verschillen. Om deze beoordeling en de eerste zorg te standaardiseren werd de *Crisis Checklist App* ontwikkeld die wereldwijd beschikbaar is voor smartphones en tablets. We vonden dat in nagebootste situaties met een ernstige zieke patiënt zorgverleners door deze app beter gingen samenwerkten en ze minder cruciale stappen van de zorg oversloegen.

In de studie beschreven in **hoofdstuk 3** wordt op de afdeling Chirurgie het gebruik van een beslissingsondersteunend systeem geëvalueerd dat zelfstandig vitale parameters kan meten bij patiënten, zoals het meten van de bloeddruk en de ademhaling. In combinatie met ingevoerde gegevens van zorgverlener berekent het systeem op basis van algoritmes hoe ziek een patiënt is en stemt het systeem zijn advies af op het protocol. In de periode dat dit systeem beschikbaar was werden er veel meer betrouwbaardere metingen geregistreerd waardoor het lokale protocol beter werd nageleefd.

In **hoofdstuk 4** wordt de innovatieve digitale checklist (TraceBook) vergeleken met papieren checklists. In tegenstelling tot de papieren checklist is de digitale checklist dynamisch en past het zijn inhoud aan op de patiënt en de lokale protocollen. De beschikbaarheid van deze checklist zorgde ervoor dat de beslissingen van de artsen tijdens dagelijkse artsenbezoek op de *Intensive Care* vaker overeen kwamen met de lokale protocollen. Dit resulteerde onder meer in een kortere verblijfsduur op de *Intensive Care*. TraceBook's digitale checklist werd beschouwd als een innovatief en educatief alternatief voor de huidige papieren checklist, maar enkele gebruikers waren

tevens bezorgd dat dergelijke beslissingsondersteunende systemen het kennisniveau van de zorgverleners kunnen aantasten ondanks de gevonden verbeteringen voor de patiënt.

De studies beschreven in **hoofdstuk 5** richten zich op de toepassing van een geautomatiseerde beademingsvorm. Hierbij regelt een beademingsmachine door middel van algoritmes zelf hoe de patiënt wordt beademd met als doel om te voldoen aan de behoeften van de patiënt en tegelijkertijd de longen zoveel mogelijk te beschermen. Uit een landelijke survey kwam naar voren dat deze beademingsvormen zeer weinig worden toegepast omdat er onder andere een gebrek is aan ondersteunend wetenschappelijk bewijs en zorgverleners angstig zijn om de controle te verliezen over het beademingsproces. Een observationele studie en een gerandomiseerde, gecontroleerde klinische studie toonden echter aan dat het toepassen van de meest geavanceerde geautomatiseerde beademingsvorm bij patiënten na een openhartoperatie leidde tot meer long beschermende beademing waarbij de patiënt sneller weer zelfstandig ging ademen.

De conclusie van dit proefschrift is dat de 4 onderzochte digitale beslissings-ondersteunende systemen alle veilig zijn en het doel behaalden waarvoor ze ontwikkeld zijn. Dit proefschrift leidt tot nieuwe aanbevelingen waaraan dergelijke systemen voor de acute en intensieve zorg moeten voldoen. Deze inzichten kunnen niet alleen gebruikt worden om de 4 onderzochte systemen te verbeteren, maar ook om zorgverleners, ontwikkelaars en onderzoekers te ondersteunen bij het ontwikkelen, testen of implementeren van beslissingsondersteunende systemen voor de zorg.

DANKWOORD

Nagenietend van een ondergaande September zon in Essendiéras besefte ik me dat "*Een visie*" het startpunt van dit proefschrift was. Een visie die resulteerde in een wetenschappelijke reis van bijna 10 jaar waarin wetenschappelijke en klinische werkzaamheden hand in hand gingen. Gedurende deze onvergetelijke reis hebben vele barmhartige Samaritanen mijn pad gekruist. Aan allen ben ik veel dank verschuldigd vanwege hun hulp, motivatie en vertrouwen. Dankzij het ontbreken van een gidsende vaderfiguur heb ik, als "The Kid", de luxe gehad om te leren van de goede eigenschappen van hen die mij intensief hebben begeleid en geholpen. Aangezien schrift gegrift waar gepraat vergaat, wil ik dit dankwoord dan ook aangrijpen om mijn dankbaarheid te vereeuwigen voor hen die mij gedurende deze reis intensief hebben bijgestaan.

Allereerst wil ik daarvoor stil staan bij mijn promotor Erik. In mijn ogen ben jij echt een visionair met 1000 en 1 ideeën die vaak hun tijd vooruit zijn: van magische pijnstillende bolletjes tot het kopen van bos, tot de eerste klinisch toepassingen van digitale klinische beslissingsondersteuning, en nu het concept: TraceBook. Jouw aankoop in Mombaroccio is dan ook geen toevalligheid gezien de affiliatie van dit Italiaanse plaatsje met Galileo Galilei. Al met al heb ik veel bewondering voor jou chaotische brein en imposante loopbaan, maar bovenal voor jou vermogen om overal ter wereld het ijs te kunnen breken. Overal weet jij een glimlach bij mensen tevoorschijn te toveren. Ernstig zieke patiënten konden weer glimlachen bij onze visites op de IC, terwijl indommelende Chinezen in een rokerig Jazz tentje ontwaakten door jou opzwepende Jazzie gepingel op de contrabas nadat we het Plein van de Hemelse vrede onveilig hadden gemaakt met jou bonte muts. Enorm bedankt voor deze levenslessen en laten we nu dit project eindigt vooral weer gaan brainstormen met een goed glas Whiskey.

Als het dan gaat over goede en kwalitatieve Whiskey, dan kan het niet anders dan dat ik nu mijn dank zal toeschrijven aan mijn copromotor Alex. Alex, tegen jou kijk ik enorm op als persoon en als jonge dokter. Jouw no-nonsens, rustige, eerlijke en betrouwbare karakter in combinatie met je grote rugzak aan medische kennis en ervaring zorgen ervoor dat ik je beschouw als één van de beste clinici die ik ken. Op de perfecte tijdstippen hebben jouw Salomonsoordelen dit proefschrift gered. Je bent mijn beste raadgever geweest. Nu al mis ik onze gesprekjes aan DE koffietafel van de IC en ik hoop dan ook van harte dat we die zullen vervangen door menig speciaal biertje in den lande. Nu dit proefschrift een feit is kan je eindelijk gaan besparen op vulpen vullingen, ongestoord cruisen met Renée in je prachtige Mercedi en vooral onbeperkt genieten van je kleinzoon en kinderen.

Mijn copromotor Arthur is de duizendpoot der duizendpotigen. Arthur, wat ben ik ongelofelijk dankbaar dat jij mij wilde helpen als lid van het promotieteam. Ik heb grote bewondering voor jou vermogen om als anesthesist de kliniek met wetenschap en het bedrijfsleven op hoog niveau te combineren naast jouw jonge gezin die bij jou duidelijk op één staat. Jouw toegankelijkheid, wetenschappelijke kennis, snelheid en excellente eigenschap om zwakke plekken aan te stippen hebben dit proefschrift naar een hoger niveau getild. Daarnaast heb ik ontelbare keren gelachen om jou droge humor en bracht je ook nog veel vertier als ware levensgenieter. Hopelijk weet je in de toekomst nog wat tijd vrij te maken om onze wetenschappelijke ambities samen waar te maken.

Bij hoge uitzondering mag ik ook nog mijn dank getuigen aan mijn copromotor Pieter. Jij bent een voorloper op het gebied van het samenbrengen van de medische wereld met de technische wetenschap en het bedrijfsleven. Jouw kennis was dan ook onontbeerlijk voor dit proefschrift. Daarnaast heb je me geleerd om zowel de kracht van lichaam als de geest te bewaken. Jouw ergonomische werkplek met moderne vernuftigheden en een hardloopsessie in Hangzhou langs Chinese bruidjes en bruidegommetjes zullen me dan ook altijd bij blijven.

Naast mijn promotieteam wil ik de leden van de promotiecommissie bestaande uit Prof. Dr Ir A.C.P.M. Backx, Prof. Dr N.J.M. van der Meer, Prof. Dr C.J. Kalkman, Prof. Dr N.H.J. Pijls, Prof. Dr Ir X. Lu, Prof. Dr Ir H. Duan hartelijk danken voor hun inhoudelijke beoordeling van dit proefschrift.

Dit proefschrift is tot stand gekomen door twee unieke en uitdagende samenwerkingsverbanden: IMPULS II (het Catharina Ziekenhuis, de Technische Universiteit Eindoven, Philips Research) en BrainBridge (de Technische Universiteit Eindoven en Zhejiang University in Hangzhou, China). Doordat ik in ieders keuken een kijkje mocht nemen ben ik enorm gegroeid als persoon en professional.

Binnen het Catharina Ziekenhuis is de intensive care het warme nest waar de visie voor dit proefschrift werd gecreëerd en waar een groot deel van mijn onderzoeken heeft plaats gevonden. Allereerst wil ik de verpleging van de IC bedanken, want ondanks mijn bombardement aan onderzoeken waren jullie altijd gemotiveerd en enthousiast om mij te helpen zelfs wanneer dit ten koste ging van jullie vrije tijd. Zonder jullie en de nodige kapsalons in de nachtdiensten was dit proefschrift nooit gelukt en was ik ook een kilo of 5 lichter geweest. In het verlengde hiervan gaat er ook veel dank uit naar Paul Kuijpers en Hans Verberne. Paul, jouw vertrouwen en leiderschap was onontbeerlijk voor al het onderzoek in dit proefschrift, terwijl Hans zijn enthousiasme over beademing en Hamilton de bakermat was voor mijn interesse in geautomatiseerde mechanische

ventilatie. Maar bovenal wil ik graag alle intensivisten bedanken voor al hun kostbare tijd en moeite. Aan jullie kan ik een heel hoofdstuk wijden. Jullie enthousiasme, motivatie voor het vak, open houding, kennis en teamwork zijn de inspiratie geweest om mijn koers van de medisch oncologie naar de intensive care te veranderen. Alex, Arnout, Herman, Erik, Dick, Marco, Marc, Femke, Ingeborg, Vesna, Luuk, Pablo, Marlijn, Joris, Harm en Leon jullie hebben aan mij echt jullie passie voor het vak door gegeven, een geweldige en onvergetelijk gift. Pas als je weg bent, besef je hoe mooi het was. Daarmee mis ik ook dagelijks onze momenten van educatie tijdens de dagelijkse visites, van koffieleuten in de ochtend en van brainstormen op jullie kamers.

Naast de IC ben ik de verpleging, alle arts-assistenten en alle stafleden van de Interne Geneeskunde dankbaar voor hun enthousiasme en deelname aan de onderzoeken die in dit proefschrift staan beschreven. Daarnaast zal ik de gastvrijheid van jullie groep nooit vergeten. Betere leermeesters had ik me niet kunnen wensen. De koffietafel levenslessen van de drie Wijzen (Alex, Arnout en Steven) zullen me altijd bij blijven. Hard werken werd afgewisseld met gezellige borrels, barbecues en vooral de grootste lol op de Oostenrijkse pistes met lange Brabantse nachten in het dal. Terugkijkende op een plezierige en leerzame tijd wil ik dan ook nog een speciaal woord richten aan mijn opleider Stijn Konings. Stijn, jij hebt mij alle ruimte en vrijheid gegeven voor dit proefschrift, naast mijn opleiding tot Internist. Ik zal jouw steun hiervoor dan ook nooit vergeten. Als internist, manager en levensgenieter ben je dan ook een van mijn grootste voorbeelden.

Beste Wilma, jij verdient een eigen paragraafje. Wat heb ik toch veel aan jou te danken als datadetective. Jij wist altijd de medische data op te sporen die ik nodig had. Zonder jou was part 4 slechts een schim geweest van wat het nu is. Daarbovenop komt dan ook nog een keer dat je een mega gaaf persoon bent om mee te lachen en te stappen, samen met onze partner in crime Britt. Laten we dat dus ook vaker gaan doen. Naast jou wil ik ook nog Harald, Tim en de de gehele afdelingen ICMT en Healthcare Intelligence bedanken voor al jullie flexibiliteit en het werk dat jullie voor dit proefschrift hebben verzet.

Als we het dan over gave personen hebben dan kom ik uiteraard aan bij mijn PhD-maatje Eveline. Voordat jij kwam was ons wetenschapshokje een grot uit de ijstijd. Altijd koud en er gebeurde nooit wat. Maar toen jij kwam fleurde alles op. Dankzij jou heb ik mijn fulltime wetenschapsjaartje overleefd. Bedankt voor al die heerlijke kopjes koffie en al je hulp, ook al kwamen mijn verzoekjes altijd last minute. Nu ik klaar ben krijg je eindelijk meer tijd voor je eigen proefschrift, maar veel belangrijker nog voor je gezinnetje met Oscar.

Ook binnen de Technische Universiteit Eindhoven ben ik aan veel mensen dank verschuldigd. Prof. Dr Ir Jan Bergmans, in mijn ogen ben jij de drijfkracht achter beide samenwerkingsverbanden en daarmee heb je achter de schermen een belangrijke rol gespeeld voor mijn proefschrift. Ik sluit me volledig aan bij jou visie dat de toekomst voor de gezondheidszorg en techniek in deze vorm van internationale samenwerkingsverbanden gezocht moet worden. Ondanks jouw imposante statuur waardeer ik je als een behulpzaam, toegankelijk en rechtvaardig persoon die altijd voor je klaar staat.

Thanks to both partnerships, I had the privilege and was fortunate to collaborate with Shan, Hui and of course Lonneke. Together with our supervisors, we all created TraceBook from scratch. I'm proud on our incredible results. I want to thank you all so much for your hard and sometimes boring work, but above all I want to thank you for all the fun we had, including our memorable trip to Hangzhou. I hope we will be able to cross swords again in the future.

Binnen de universiteit gaat ook nog een woord van dank uit naar Prof. Dr Ir Uzay Kaymak. Jou adviezen hebben me veel geholpen en je gaf mij de eer voor mijn eerste lezing op een universiteit. Ondanks dat ik als arts toch een vreemde eend in de bijt ben van een Technische Universiteit voelde ik me altijd direct thuis dankzij de vele koffietjes, borrels en diners met mijn TuE-maatjes Tilaï, Ineke en Dennis. Bedankt voor al jullie adviezen en brainstormsessies.

Naast de Universiteit was Philips Research op de High Tech Campus een indrukwekkende omgeving die mij ontzettend veel inspiratie en motivatie heeft geboden voor mijn proefschrift. Dankzij de expertise van Jacco Eerden, Jelle Tuinhout, Alexander Heuts, Gertjan Maas en Jasper Stam heeft TraceBook de stap gezet van een prototype naar een modern beslissingsondersteunend systeem dat klinisch getest kon worden. Daarvoor ben ik Franklin Schuling en Steffen van Leeuwen ook enorm dankbaar want zonder hun coördinatie, motivatie en vertrouwen was deze metamorfose nooit tot stand gekomen. Daarnaast hebben Lenneke van Genugten en Kiran Dellimore het real-life onderzoek van TraceBook naar een hoger niveau getild. Jullie deskundigheid bood de onontbeerlijk details en diepgang om daarmee de implicaties van de gevonden kwantitatieve resultaten beter te doorgronden voor deze nieuwe technologie. Wie echt niet mag ontbreken in dit lijstje van belangrijke mensen voor mij binnen Philips Research is Rick Bezemer. Jouw werklust, intelligentie en sociale skills als extreme motivator hebben me geregeld aangezet om onbereikbaar geachte dromen en doelen na te streven. Dit proefschrift is daar het resultaat van en ik hoop dan ook van harte dat we in de toekomst nog mogen samenwerken en bieren!

Dear Marcus Schultz and Ary Serpa Neto, you've taught me that science is a form of art. While working with you for one day, I learned more skills about performing research and writing scientific manuscripts than I ever learned from reading books and the internet. I'm eternally grateful for all your patience, wise lessons, revisions, and kindness.

I'd also like to sincerely express my appreciation to Chris Subbe, John Kellett, Paul Barach, Tim Cooksley, and Mikkel Brabrand. Thank you for your precious time and all your assistance with the scientific work concerning the Crisis Checklist App and my thesis. I'm looking forward to have a drink and celebrate all the milestones of our collaboration.

Ontspanning en inspiratie zijn essentieel voor een proefschrift. Daarmee wil ik ook mijn mede-Fellows van de IC in het RadboudUMC bedanken voor hun flexibiliteit als ik weer eens wetenschappelijke deadlines had of voor mijn gechagrijn als ik weer een wetenschappelijk hobbel tegen kwam. Aart, de rust zelve en een geweldig klankbord. Helaas kennen we elkaar nog maar kort, maar onwetend heb je mij echt door deze drukke tijd heen geloodst. Je was mijn bierdrinkmaatje die samen met Stephanie mij veel koppijn heeft bezorgd in Brussel. Laten we nog vaak toosten op onze vriendschap en PSV. Jessica, mijn IC-yin, wat heb ik genoten van onze wetenschappelijke en klinische discussies. Ik weet zeker dat je het met me eens bent dat dit proefschrift voer is voor de kanonnen. Daarnaast wil ik nog aan Wouter mijn excuses bieden voor het ontbreken van een NaBic-regel in TraceBook. Naast de IC-Fellows wil ik alle stafleden van de IC in het RadboudUMC en in het bijzonder mijn opleider IC Prof. Dr Hans van der Hoeven bedanken voor al hun steun, gegeven ruimte en vertrouwen in mij. Jullie hebben me met open armen ontvangen en enorm veel kennis op het gebied van de IC gegeven. Hans, zonder jou had ik al mijn functies als fellow, promovendus en jonge vader nooit kunnen combineren. Je bent een geweldig voorbeeld voor mij als wetenschapper, als dokter en als oprecht en goedhartig mens.

Mannen van Trega, wat hebben jullie me onvergetelijke tijden bezorgd waardoor ik het leven als medicus en wetenschapper telkens weer kon loslaten voor broodnodige ontspanning. Jullie hebben me altijd gesteund met dit gekkenwerk, vanuit Krakau tot in Alicante. In het bijzonder wil ik ook de binken Mickey, Serve, Erwin, Casper, Jouke, Dirk, Mark, Ruben, Rutger en Ronnie bedanken voor hun steun tot in de vroegste uurtjes.

Bart en Diederick, ook jullie mogen niet ontbreken in dit dankwoord. Roomie, zonder onze uitstapjes en filosofische discussies was ik nooit man geweest die ik nu ben. Jouw nuchterheid en kijk op het leven hebben mij en dit proefschrift heel veel gebracht. Nu maar hopen dat ik met mijn vrije tijd wat beter kan leren schaken om je te verslaan. Diederick, mogelijk besef je het van jezelf niet, maar je bent een van de grootste

motivators die ik ken. Een echte trouwe vriend die mij binnen een dag weer kan laten geloven in mijn dromen en ambities, ondanks dat je me dan sportief keer op keer weer keihard inmaakt.

Beste Paranimfen, Stephan en Frederik, wat kijk ik tegen jullie op. Het schrijven van dit proefschrift stelt eigenlijk niks voor met jullie daden als jonge drukke vaders met ieder hun tweeling. Stephan, jouw trouwe eerlijke karakter, "down to earth" mentaliteit, en humor hebben me door jaren van zwoegen geholpen. Je hebt me de wijze les geleerd dat het nimmer in de eerste plaats om jezelf moet gaan, maar dat je gezin op één staat. Je zorgde altijd weer voor de perfecte afleiding door lekker over sport te leuteren, dus laten we met de komst van meer vrije tijd ook een datum prikken om samen een groot sportevenement live mee te maken. Frederik, als Freshley gaan we nog steeds samen door het leven en wat geniet ik daarvan. Veel plannen en wetenschappelijke uitwerkingen in dit proefschrift zijn een uitkomst van onze discussies. Er is niks zo chill en enthousiasmerend om met jou te borrelen en te zwetsen over het leven. Je bent zeker de slimmere van ons twee en dat heeft me altijd extreem gemotiveerd om je proberen te evenaren. Achterlopen en falen was dan ook voor Freshley geen optie. Ook jouw opleiding en promotietraject beginnen nu tegen het einde te lopen en ik weet dan ook zeker dat we daarna de tijd zullen vinden om die Mont Blanc samen te beklimmen.

Lieve Pappy en Mammy, dankzij jullie liefde en geborgenheid ben ik op dit punt gekomen. Pappy, streng edoch rechtvaardig heb je mij meegegeven dat alles wat moeite kost de moeite waard is. Iedere keer weer en nu ook met dit proefschrift heb je weer helemaal gelijk. Mammy, de mater familias en mijn meter, voor jou kon ik niks verkeerd doen. Jouw pompoensoep was onevenaarbaar. Jouw glimlach was en is er nog altijd. Pappy en Mammy, ik weet dat jullie vanuit de hemel met trots naar mij glimlachen.

Lieve Jaap en Hanny, in deze tropenjaren waardeer ik jullie steun aan Marijke en mij heel erg. Zonder jullie hulp hadden we onze proefschriften nooit af kunnen maken. En wat is een gepromoveerde vader nou zonder de basis, van professioneel barbecueën tot fietsenbanden plakken en het vervangen van autobanden. Jaap, ontzettend bedankt voor al je geduld gezien mijn twee linker handen. Michiel, jij verdient ook een speciaal woordje aan dank want je staat altijd voor ons klaar en hebt zelfs met je WESP een belangrijke bijdrage geleverd aan dit proefschrift. Beppe, haal de Rummikub maar boven, want ik ga dit spel als een monnik bestuderen om meer tegenstand te kunnen geven nu dit proefschrift tot een einde is gekomen.

Lieve diva's van mijn leven, lieve Mama en Sissies, met z'n vieren zijn we door zeer turbulente jaren gekomen die mij heel sterk hebben gemaakt. Mamma, als

alleenstaande vrouw heb je me laten zien hoeveel je kan bereiken. Jouw veerkracht, doorzettingsvermogen, wijsheid, geduld en vertrouwen in ons hebben ertoe geleid dat Isabel, Barbara en ik zo ver zijn gekomen met allen een universitair diploma en nu ook dit proefschrift. Prachtig hoe je me hebt geleerd om hard te werken en in je dromen te geloven. Met je nieuwe appartement in ODK heb je met Wim dan ook weer een wens in vervulling laten gaan en begin je steeds meer op Mammy te gelijken. Daar gaan je kleinkinderen en mijn maag nog heerlijk van genieten. Isabel en Barbara, jullie hebben me vooral veel tijd gegeven. Op de stoep voor de Zara, Victoria Secret en de Bijenkorf had ik zeeën van de tijd om met kledingzakken in beide handen na te kunnen denken over dit proefschrift. Jullie waren altijd mijn kleine zusjes, maar het is geweldig om jullie nu zo sterk op je eigen benen te zien staan. Tot dit proefschrift had ik me ook nooit gerealiseerd hoe handig het is om twee zusjes te hebben met zoveel communicatief talent. Een vloeiende publiekssamenvatting is het resultaat. Dank jullie voor alles en laten we samen nog veel lol beleven.

Lieve Louise en Nina, als jullie dit proefschrift gaan lezen zullen jullie lachen om de prehistorische gadgets die jullie papa onderzocht. Gedurende mijn wetenschappelijke reis hebben jullie mij de belangrijkste levensles geleerd. Sommige emoties kende ik niet goed of wist ik niet meer hoe ze voelde voordat jullie in ons leven kwamen. Jullie hebben mij geleerd hoe één glimlachje een sombere zware dag kan openbreken, wat angst is als jullie even uit het zicht zijn, hoe medelijden voelt als jullie je bezeren, hoe goed je conditie moet zijn om papa Gruffalo te zijn, hoe het voelt om een held te zijn bij het plakken van een pleister, hoe mooi het is om pirouetjes te draaien op Disney muziek, maar bovenal hoe ontzettend veel je van twee prachtige prinsesjes kan houden. Dankzij jullie kan ik mijn verstand op nul zetten en genieten van de kleine dingen des levens met al onze mooie avonturen.

Aller liefste prachtige Marijke, vanaf het moment dat ik je zag op de driehoekige berg van de universiteit Maastricht geloof ik in het lot. We zijn een perfect team. We halen het beste bij elkaar naar boven, ook al bereik jij de mijlpalen, uiteraard, een fractie eerder. Ondanks je drukke leven als promovendus, internist i.o., kersverse nefroloog en als de allerliefste moeder neem je altijd de tijd voor mij. Ik zal dan ook mijn best doen om mijn gehate uitspraakje “nog eventjes” vanaf nu drastische te verminderen. Onze prachtige reizen naar Nieuw-Zeeland, New York, Lapland en Schotland gaven mij de onontbeerlijke energie om door te gaan, terwijl jouw eerlijkheid, wetenschappelijke ervaring en wijsheid ook mijn proefschrift naar een hoger niveau hebben getild. Van alle lof die aan dit proefschrift wordt toebedeeld, behoort dan ook een groot deel aan jou toe. Zonder jou was dit proefschrift er namelijk nooit geweest. Maar bovenal, zonder jou was ik nooit zo gelukkig geweest.

CURRICULUM VITAE

Ashley Jacobus Raphaël De Bie was born on November 7th, 1986 in Eindhoven, The Netherlands. In 2005, he graduated from secondary school at Gymnasium Beekvliet, in Sint-Michielsgestel, and entered Maastricht University to study medicine. In 2011, he graduated with a master of sience degree in medicine. Following short he moved to Eindhoven and started as a physician in the Intensive Care Unit of the Catharina Hospital in Eindhoven, the Netherlands. After a year he made a career transition to the Internal Medicine of the Catharina Hospital where he started his residency training in 2014 under the supervision of Dr C.J.A.M. Konings. Between 2014 and 2020, he carried out his doctoral research under the supervision of Prof. Dr H.H.M. Korsten, Dr A.J.G.H. Bindels, Dr R.A. Bouwman, and Dr P.M.E. Van Gorp in a collaborative framework between the Catharina Hospital, the Eindhoven University of Technology, and Philips Research – as part of the IMPULS-II program. Since 2019 he has been employed as a fellow in the Intensive Care Unit of the Radboud UMC to follow a specialized training in Intensive Care Medicine under the supervision of Prof. Dr J.G. van der Hoeven. Ashley is married to Marijke Dekker since August 2014 and both are the proud parents of two daughters, Louise (2017) and Nina (2019).

PUBLICATIELIJST

- **Fully automated closed-loop ventilation is safe and effective in post-cardiac surgery patients.** De Bie A.J.R., Roos A.N., Bindels A.J.G.H. *Intensive Care Med* 2014; **40**: 752-3

- **Thinking beyond the mass: ANCA-associated vasculitis mimicking a pancreatic malignancy.** De Bie A.J.R., Dekker M.J.E., Vermeulen Windsant I.C., Nikkessen S., Demeyere T.B., Konings C.J., de Hingh I.H., Franssen C.F., Creemers G.J. *Neth J Med* 2015; **73**: 341-4

- **The correlation of age with chemotherapy-induced ovarian function failure in breast cancer patients.** Vriens I.J., De Bie A.J.R., Aarts M.J., de Boer M., van Hellemond I.E., Roijen J.H., van Golde R.J., Voogd A.C., Tjan-Heijnen V.C. *Oncotarget* 2017; **8**: 11372-9

- **Current practice of closed-loop mechanical ventilation modes on intensive care units - A nationwide survey in the Netherlands.** De Bie A.J.R., Wenstedt E.F.E., Roos A.N., Verberne J.J.M., Korsten H.H.M., Schultz M.J., Bindels A.J.G.H. *Netherlands Journal of Medicine* 2017; **75**: 145-50

- **Intelligent dynamic clinical checklists improved checklist compliance in the intensive care unit.** De Bie A.J.R., Nan S., Vermeulen L.R.E., Van Gorp P.M.E., Bouwman R.A., Bindels A., Korsten H.H.M. *Br J Anaesth* 2017; **119**: 231-8

- **Crisis checklists for in-hospital emergencies: expert consensus, simulation testing and recommendations for a template determined by a multi-institutional and multi-disciplinary learning collaborative.** Subbe C.P., Kellett J., Barach P., Chaloner C., Cleaver H., Cooksley T., Korsten E., Croke E., Davis E., De Bie A.J.R., Durham L., Hancock C., Hartin J., Savijn T., Welch J., Crisis Checklist C. *BMC Health Serv Res* 2017; **17**: 334

- **An alternative ICU staffing model: implementation of the non-physician provider.** Kreeftenberg H.G., Aarts J.T., De Bie A.J.R., Bindels A.J.G.H., Roos A.N., van der Voort P.H.J. *Neth J Med* 2018; **76**: 176-83

- **A paraplegic patient with fever and leucocytosis: not always what it seems.** De Bie A.J.R., S.K., Kersten E., Kisters J.M.H., Vrijhof H.J.E.J., Willems J.. *A paraplegic patient with fever and leucocytosis: not always what it seems* 2019; **15**

- **Impact of the Advanced Practice Provider in Adult Critical Care: A Systematic Review and Meta-Analysis.** Kreeftenberg H.G., Pouwels S., Bindels A.J.G.H., de Bie A.J.R., van der Voort P.H.J. *Crit Care Med* 2019; **47**: 722-30

- **Implementation of an automated early warning scoring system in a surgical ward: Practical use and effects on patient outcomes.** De Bie A.J.R., Mestrom E., Steeg M.V., Driessen M., Atallah L., Bezemer R., Bouwman R.A., Korsten H.H.M. *PLoS One* 2019; **14**: e0213402

- **Differences in identification of patients' deterioration may hamper the success of clinical escalation protocols.** De Bie A.J.R., Subbe C.P., Bezemer R., Cooksley T., Kellett J.G., Holland M., Bouwman A.R.A., Bindels A., Korsten H.H.M. *QJM* 2019

- **Identify facilitators and challenges in computerized checklist implementation.** Nan S., De Bie A.J.R., Zhang S., Korsten H.H.M., Lu X., Duan H. *Stud Health Technol Inform* 2019; **264**: 1737-8

- **Post-ICU care after a long intensive care admission: a Dutch inventory study.** Hendriks M.M.C., Janssen F.A.M., te Pas M.E., Kox I.H.J.M., van de Berg P.J.E.J., Buise M.P., De Bie A.J.R. *Netherlands Journal of Critical Care* 2019; **27**: 190-5

- **Fully automated postoperative ventilation in cardiac surgery patients: a randomised clinical trial.** De Bie A.J.R., Serpa Neto A., van Meenen D.M., Bouwman R.A., Roos A.N., Lameijer J.R., Korsten H.H.M., Schultz M.J., Bindels A.J.G.H. *Br J Anaesth* 2020; **125**: 739-49

- **Testing the effects of checklists on team behaviour during emergencies on general wards: An observational study using high-fidelity simulation.** De Bie A.J.R., Dijkmans J.J., Todorovac N., Hibbs R., Boe Krarup K., Bouwman R.A., Barach P., Fløjstrup M., Cooksley T., Kellett J., Bindels A.J.G.H., Kortsten H.H.M., Brabrand M., Subbe C.P. *Resuscitation* 2020

- **Intelligent checklists improve checklist compliance in the intensive care unit: a prospective before-and-after mixed-method study.** De Bie A.J.R., Mestrom E., Compagner W., Nan S., van Genugten L., Dellimore K., Eerden J., van Leeuwen S., van de Pol H., Schuling F., Lu X., Bindels A.J.G.H., Bouwman R.A., Kortsten H.H.M. *Br J Anaesth* 2020