Can simulation based training for emergency response teams improve mortality rates, reduce adverse events and risk of errors?

– A systematic literature review.

Faculty of Health Sciences

Master i Pre-Hospital Critical Care

Master Thesis (30 ECTS)

Name of student – Kjetil Torgeirsen

Name of supervisor – Thomas W. Lindner

02.06.2020

UNIVERSITY OF STAVANGER

MASTER IN PRE-HOSPITAL CRITICAL CARE

Master thesis

Semester:

Spring – 2020

Author/Thesis candidate: Kjetil Torgeirsen

Supervisor: Thomas W. Lindner

Thesis Title:

Nordisk tittel (Norwegian title):

Kan simuleringsbasert trening for akuttmedisinske beredskaps team – øke overlevelse og redusere antall uønskede hendelser og risiko for behandlingsrelaterte feil? En systematisk gjennomgang av litteratur.

English title:

Can simulation based training for health care emergency response teams improve mortality rates, reduce number of adverse events and risk of errors? A systematic literature review.

KEYWORDS/SEARCH WORDS:

Simulation, emergency response teams, improved patient outcome

---

NUMBER OF PAGES: 59 + attatchments

STAVANGER: 2. juni 2020

DATE/YEAR

Preface

This thesis marks a milestone in an academic journey that started in august 2016. It is written as part of the Pre-hospital critical care master's program at the University of Stavanger. The work with this thesis is a summary of my professional career so far. My 25-year clinical background from the Stavanger University Hospital EMS and 14 years working with medical simulation had a strong influence on the choice of topic for the thesis. I have always had a strong interest in the quality of care in pre-hospital emergency care and my ambition for the master thesis was to contribute to the professional fields of emergency medicine and medical simulation.

I want to thank my supervisor, Thomas Lindner, for helpful discussions and the invaluable efforts he has put into supervising the process of writing this thesis. I also want to thank Corinna Vossius and Patrik Nystrøm for their contributions. Corinna provided great additional advice and guidance during the last phase of the writing process. Patrik was the third reviewer in the review process and provided feedback through discussions during the writing process. Thank you all three for your contributions and collaboration.

Thanks to all past and present colleagues from 25 years in the ambulance service and to my SAFER colleagues. A personal thank you goes to Liv Norland for her technical support with conversion and merging of documents. Thank you to librarian Hilde E.S. Magnussen for your help with search strategies and literature searches.

I also want to thank my family for their support during the 4 years I have attended the PHCC master's program. A special thanks to my wife Solveig for her patience and for covering "the home base" in times of absence.

Table of contents

Tables

Figures

Summary

Background:

The objective of this master thesis was to examine the following research question: Can simulation-based training for health care emergency response teams reduce mortality rates, number of adverse events and risk of errors?

Methods:

A systematic literature review based on the PRISMA guidelines was conducted in the following databases: Medline, EMBASE, Epistemonikos, Google Scholar, PubMed, CINAHL and Cochrane. Three reviewers used pre-defined criteria and the Rayyan QRCI© tool to screen the retrieved studies independently for inclusion in the analyses. From those studies the results of simulation-based training according to Kirkpatrick level 4 where extracted. A tool for evaluation of quality of reporting the simulation-based training intervention was developed and applied.

Results:

184 studies have been identified and 33 were further investigated. On the Kirkpatrick level 4 thirty studies (90%) described improvements, while three (9%) found no changes. 19(58%) of the studies were rated 4/7 points or higher for the quality of reporting of the simulation- based training intervention. 18 studies (55%) reported repeated simulation training as part of intervention and reported reduced mortality rates and reduced numbers of adverse events. None of the included studies reported numbers of medical errors.

Conclusion:

The results indicated that simulation based training can reduce mortality and adverse events. Repeated training seemed to increase the possibility for Kirkpatrick level 4 improvements. This systematic review was not able to identify other characteristics of simulation based training that affected patient outcome.

1.0 Introduction

In this chapter, I will explain my professional background, the background and objectives for the thesis.

I have a background as a paramedic working in prehospital emergency medical services (EMS) for 25 years. I have also worked at the SAFER simulation centre since 2006. The work with medical simulation at SAFER has given me a personal experience of the clinical potential of simulation as a training modality and as an efficient tool for transfer of knowledge and skills into clinical practice. My background and my interest in emergency medicine and simulation as a training modality has led to my participation in the Pre-Hospital Critical Care program at the University of Stavanger and to the topic for this thesis.

The thesis has been written according to the PRISMA checklist. Some adaptions had to be made due to the MPHMAS thesis guideline, e.g. the theory chapter which is not a part of the PRISMA checklist.

1.1 Rationale

The publication of the report "To err is human" in 1999 (1) may be considered as a starting point of patient safety. The conclusion from this report was that medical errors lead to a high number of preventable deaths in hospitalized patients, and these findings are frequently quoted: "Over a decade ago, The Institute of Medicine disclosed that error is a significant cause of death in the United States that accounts for 44,000 to potentially as many as 98,000 deaths annually. This magnitude makes errors more lethal than motor vehicle collisions, breast cancer, and AIDS." (2)

More recently, the Worlds Health Organization (WHO) has initiated the campaigns "Patient Safety - Making healthcare safer" (3) from 2017 and "10 facts on patient safety" (4) from 2019. Both campaigns focus on patient safety and implications from unsafe patient care. Some of the evidence-based facts presented in these campaigns are: Patient safety is a serious global public health problem. It is estimated that 1 of 10 patients in high income countries is harmed while receiving hospital care. Harm might be caused by a range of adverse events where 50% are considered to be preventable. Adverse events trough unsafe care is estimated to occur in 42,7 million hospitalizations and are likely to be one of the ten leading causes of death and disability globally.

The need for tools to improve quality of care and patient safety in health care have been described in several reports and articles, and simulation, defined as “a model or mock-up for purposes of experiment or training” (5) was introduced as a tool to increase patient safety and reduce number of human errors (1, 2, 6)

1.2 Objectives and aims

The objective of this master thesis is to examine the research question: Can simulation based training (SBT) for health care emergency response teams reduce mortality rates, number of adverse events and risk of errors?

The primary aim is to investigate to what degree SBT for emergency medical teams has an impact on patient outcome defined as Kirkpatrick level 4 (K4 level) (7) measured as mortality rates, number of adverse events and medical errors by applying a systematic literature review methodology to summarize recently published research.

In addition, it is to investigate a) the quality in reporting the applied SBT interventions, b) if there were any SBT design similarities in studies related to reported results and c) whether study design affected sustainability of changes.

1.3 Relevance of the project

The recognition of the scale of medical errors and patient safety problems in health care was the first step in learning about common causes of errors and searching for efficient methods to improve patient safety and prevent medical errors, as Charles Vincent (8) states: ”There is compelling evidence that, while healthcare brings enormous benefits to us all, errors are common and patients are frequently harmed.”

The WHO (4) has declared adverse events as one of ten leading causes of deaths globally. Compared to other high-risk industries, the risk of patient death caused by a medical accident healthcare is much higher. For aviation the likelihood of dying is estimated to be 1 in 3 million, while the likelihood for dying due to a medical accident in hospitals is 1 in 300.

SBT as a training modality is believed to be efficient for students, novices and experienced clinicians (2, 6, 9) . The WHO (10) has stated simulation to be an important component in the work of improving patient safety. In Norway, the National health and hospital plan 2020-2023

(11) clearly states the importance of simulation to ensure adequate and correct expertise to enable the health care system to meet future challenges and needs. Despite the recognition of SBT from both medical educational institutions and official health care organizations, past systematic reviews have found conflicting results regarding the effect of SBT on patient outcome, and there is no given description of the characteristics training should have in the official recommendations. It is also unclear how different SBT designs affects efficiency and how to design high quality and efficient SBT interventions that could be expected to improve patient safety and patient outcome. There is a need for more research searching for the causes of conflicting results and to investigate the evidence for K4 level effects of SBT and how quality of SBT and different SBT designs influences the efficacy and efficiency of the SBT intervention.

1.4 Previous research:

In 2019 scoping review of the existing evidence of SBT and K4 level effects was performed. The review was written as a part of the patient safety module in the PHCC master's program at the University of Stavanger using the PRISMA framework. The review was presented as an abstract at the annual European simulation conference (SESAM) and later published in Advances in Simulation (12). Despite significant limitations, it revealed the following findings:

1. Twelve single studies reported results indicating a possible relationship between SBT and improved patient outcomes and reported of reduced numbers of adverse events (13-24).
2. Five systematic literature reviews reported mixed/conflicting results on the K4 level effects. Some studies that provided significant results were missed by all these reviews. In addition, new and important research has been published after these systematic literature reviews were performed (25-29). A new review including recent research was needed.
3. Since some of the included studies in the scoping review were performed in low resource health care systems in developing countries, the scoping review discussed whether results from studies performed in low resource health care systems could be generalizable in a high resource health care system. The scoping review found several studies from both high- and low resource health care systems presenting very similar improvements of patient outcome after SBT interventions (14, 19-21).

4. A common and recurring problem for these literature reviews was the lack of adherence to reporting guidelines for simulation based research as described by Cheng et al. (30).

A search for review articles missed by or published after the scoping review identified several reviews investigating the impact from SBT on patient outcome, presenting positive findings, but focusing on narrow fields of medicine e.g. obstetrics or paediatrics, the Helping Babies Breathe bundle, obstetric emergency care, in-situ simulation trainings, training for nurses etc. (28, 31-34). None of the identified reviews included interprofessional emergency team training across different fields of medicine. Quality assessment and analyses comparing different simulation designs and whether they have effects on patient outcome, were not performed, and conflicting results of single studies were not further investigated by other researchers (31-34).

1.5 Context

1.5.1 Medical emergency situations

The characteristics for a medical emergency, regardless of the cause, is the sudden onset, time-criticality, a need for immediate treatment, and the high stakes involved for both patient and health care providers (HCP) due to risk of disability or death for the patient, and risk of doing harm or medical errors for the HCP. These factors might add to elevated levels of stress for health care providers. Flin et al.(35) explains stress in this context to be ”emergency stress or critical incident stress” with a sudden onset, “being intense and of relatively short duration.” The main issue for the HCP is the disruption of goal oriented behaviour triggering a “fight, flight or freeze response.” Stress can be caused by the situation itself and co-factors like cognitive workload, teamwork factors, communication, awareness, fatigue, leadership issues that further increase stress levels in a situation where the HCP needs to perform at the peak of the capacity. Stress has been linked to accidents and errors and should be recognized and managed both at an individual and team level (35)

To enable HCP to perform in situations like this there are several training strategies that might be useful:

- Automation of skills to reduce cognitive workload.

- Increase of self-efficacy through training of realistic and relevant emergency scenarios.
- SBT with scenarios where successes can be reproduced in similar real-life scenarios.
- Training in the use of cognitive tools e.g. timeout, SBAR (situation, background, assessment, recommendation) closed loop communication

1.5.2 Medical emergency teams

A Medical emergency team (MET), is a team of medical professionals from different specialities trained to respond rapidly to suspected medical emergencies with the aim to identify threats to or failing of vital patient functions and to initiate treatment to prevent development or correction of a life threatening condition for the patient, e.g. trauma, cardiac arrest, paediatric emergencies and obstetric emergencies (36). Specifics of such teams are a) they are interprofessional, b) the persons in these teams are not necessarily the same every time and c) the role is connected to the speciality needed in the team and will be covered by the person on duty.

1.5.3 Simulation

Simulation has rapidly grown in popularity and scientific evidence from different medical specialities is rapidly increasing. Some specialities operate in physical environments that make it easier to perform research on simulation interventions, while other specialities are very complex with patients moving between various wards making it very difficult to point to the source of improved patient outcome.

1.6 Abbreviations

ER – Emergency room

EMS – emergency medical system

GRADE – Grading of Recommendations, Assessment, Development and Evaluations

HCP – Healthcare provider and healthcare providers

STROBE – Strengthening The Reporting of Observational Studies in Epidemiology

CONSORT – Consolidated Standards of Reporting Trials

PHCC – prehospital critical care

SBAR –Situation – Background – Assessment - Recommendation

MET – Medical emergency team, medical emergency teams

RRT- Rapid response team, rapid response teams

RRS – Rapid response system

SBT – Simulation based training

WHO – World Health Organization

2.0 Theory

2.1 Theoretical perspective

The theory chapter presents theories relevant for understanding SBT and concepts related to it, the discussion of the results from the literature search performed, and of the findings in the included studies.

2.2 Simulation as a training modality.

The use of patient actors or simulated patients in medical education was introduced in 1963 (37). Computerized simulation was introduced in the 70's and simulation was introduced as a training modality in the late 1980's. The Society in Europe for Simulation Applied to Medicine (SESAM) was established in 1993 (38).

There are several different definitions of simulation training. In the simulation dictionary (5) , a publication from the simulation community, we find six different definitions. "A technique that creates a situation or environment to allow persons to experience a representation of a real event for the purpose of practice, learning, evaluation, testing, or to gain understanding of systems or human actions" is a consensus definition, while others evolve from pioneering articles. Examples from these definitions are:

"A strategy in which a particular set of conditions are created or replicated to resemble authentic situations that are possible in real life. Simulation can incorporate one or more modalities to promote, improve, or validate a participant's performance" (6),

"A pedagogy using one or more typologies to promote, improve, or validate a participant's progression from novice to expert" (39).

Most importantly, simulation is more than an accidental scenario-based training session. There have been several attempts to characterize high quality simulation, in this context meaning to enhance the probability of clinical impact and of positive patient outcome.

A wide variety of simulation methods and instructional design features exists, but a clear definition of how simulation should be structured is lacking.

There have been some initiatives to describe standards of best practice for simulation. The International Nursing Association for Clinical Simulation (40) published a list of criterions with the aim of guiding effective simulation training:

1. Needs assessment
2. Measurable learning objectives
3. Format of simulation
4. Clinical scenario or case
5. Fidelity
6. Facilitator/facilitative approach
7. Briefing
8. Debriefing and /or feedback
9. Evaluation
10. Participant preparation

Cheng et al.(30) presented characteristics for high quality simulation based on expert consensus in the following dimensions:

1. Description of scenario and learning objectives
2. Instructional design; duration and frequency/repetitions
3. Feedback/debriefing; structure/method, duration, facilitator characteristics
4. Time used for intervention, repetitions, education of facilitators etc.

Cheng et al. also points out the importance of a description of the intervention in any study involving simulation. This will be described in more detail under 2.4.

SBT is considered a useful tool to address the gap between knowing and doing and as Flanagan et al. (9) described, also an effective tool to train human factors skills. Gaba (6) presented in his 2004 article *The future vision of simulation in health care* that SBT could be used in different contexts, for different participants and that it could improve patient safety if

used as intended. He stated that the people working within the simulation community expected a future revolution in health care with SBT as an enabling key. The article also described health care education to focus mainly on basic scientific education, emphasizing individual knowledge and skills rather than focusing on training the performance of medical teams. The aim of Gaba's article was to "provide a comprehensive framework for understanding the diversity of applications of SBT in healthcare." To enable this, he categorized the simulation application by 11 dimensions and demonstrated how SBT can be used for students as well as experienced clinicians:

"**Dimension 1: The purpose and aims of the simulation activity"**
According to Gaba, simulation could be used for different activities: Training, education, assessment of performance, a "bottom up" tool for enhancing safety culture in health care, system probing, testing of new infrastructure to identify gaps in new hospital environments etc.

"Dimension 2: The unit of participation in the simulation"
SBT could be used for individual training for teaching of knowledge and skills. SBT can also be applied to train human factor skills in teamwork with focus on tools like crew resource management (CRM) etc. It is an excellent way of training the team as a collaborating entity, but also for team members to familiarize with, and exercise their role, learning about other roles in the team and the assigned tasks for each role.

"Dimension 3: The experience level of simulation participants
Gaba states that "simulation can be applied from cradle to grave." Hence it needs to be designed according to the expected participants knowledge and level of experience to maximize the potential learning outcome. Participants should already possess the most critical knowledge, skills and experience required to solve the problems and challenges faced during the simulation. Experiences from our simulation center have shown that if participants are presented to a situation way beyond the individual or team capacities, there is a risk of frustration and other emotions that might limit the potential learning from the scenario.

"Dimension 4: The health care domain in which the simulation is applied"
SBT is applicable to most health care domains. It is important that the simulation activity includes relevant medical skills and procedures and that they can be fully and realistically performed. That the SBT itself is perceived to be relevant and that the environment is

realistic. We do not need to replicate every detail in a ward or a theatre, but realism is believed to add to the transferability of the SBT experience into clinical practice (41).

"Dimension 5: The health care disciplines of personnel participating in the simulation"

SBT is a useful training method for most health care professions. As stated before, it is believed to be an efficient tool for training of interdisciplinary teams also during their educational programs. If students learn about other professions, their assets and limitations during education, they might have a better chance of collaborating in an efficient way in interdisciplinary teams in collaboration with more experienced team members as novice HCP in clinical work.

"Dimension 6: The type of knowledge, skill, attitudes or behavior addressed in simulation"

SBT can be used to challenge, modify and make participants aware of attitudes in an unpreceded way. However, experiences from the simulation community is that this require skilled and trained facilitators to facilitate debriefing of sensitive topics e.g. attitudes in a constructive manner (42). SBT can also be used to train performance of skills and procedures in a clinical context with or without some stressors applied to the context. It is also an opportunity for observation of other participants approach to the same scenario.

"Dimension 7: The age of the patient being simulated"

SBT must use appropriate equipment (simulators) or standardized patients to make the presentation of the patient realistic. Sometimes, the scenario has to be adjusted according to age and gender of available equipment or standardized patient.

"Dimension 8: The technology applicable or required for simulations"

What kind of technology is needed to achieve learning goals? If the goal is to train emergency airway procedures, the requirement will be of a mannequin allowing the relevant procedures to be performed. If resources are limited, adjustment of learning goals could be the solution and key to successful SBT e.g. to change learning objectives from advanced airway procedures, to an ethical challenge or human factors e.g. use of communication tools if no advanced simulator allowing for advanced airway procedures are available.

"Dimension 9: The site of simulation participation"

In-situ SBT have gained popularity. It is an efficient way of conducting SBT in realistic environments known to the participants that could make SBT more "available" as an everyday training modality. As an additional benefit, the SBT can be used for system probing,

identifying gaps that need to be addressed and handled. But there is a risk of participants being distracted by real patients or clinical work. The use of equipment might also be a challenge. Sometimes a dedicated simulation center is better suited to allow participants use e.g. highly sophisticated simulators, video-recordings of the scenario or if repeated series of simulations are a part of the SBT design.

"Dimension 10: The extent of direct participation in simulation"

SBT should not be restricted to simulation centers. As the WHO (3) have pointed to the global challenges for patient safety there is a need to explore technological tools allowing remote access to simulation experts and SBT.

"Dimension 11: The feedback method accompanying simulation"

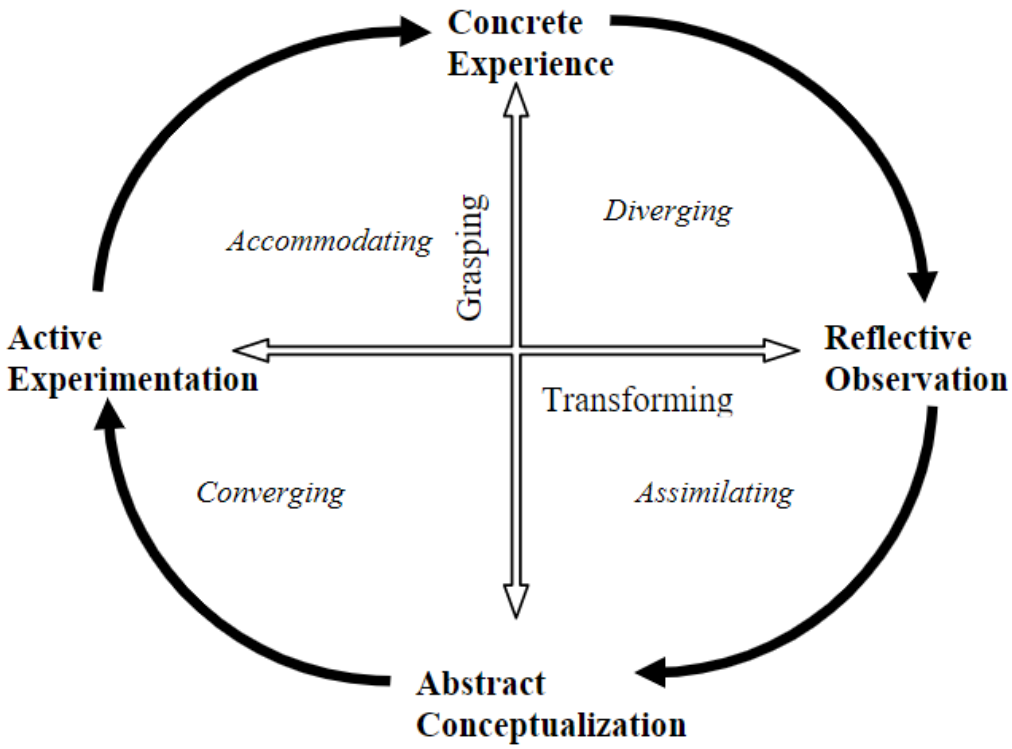
Feedback for complex SBT is commonly provided by human instructors or facilitators. The instructors should be trained in debriefing techniques to be able to facilitate participants reflection and discussions of the behaviors, decisions and actions during the scenario (42). Since SBT is exposing participants to highly realistic clinical situations the SBT is emotional potent and gaps needs to be addressed in a nonjudgmental manner to enable learning.

2.3 Pedagogical theories and simulation

To be able to understand the complexity of the learning processes involved in SBT, it is important to know something about the pedagogical theories which form the foundation of SBT as a training modality. In a review article from 2017 Anna Abelson (43) identifies Kolb and Dewey as the two major influences on simulation pedagogics.

Kolb's theory on experimental learning introduced in 1983 reflection based learning, where students/participants reflect on performance in a simulated or real scenario (44, 45).

Figure 1. Kolb's circle of learning.



(44)

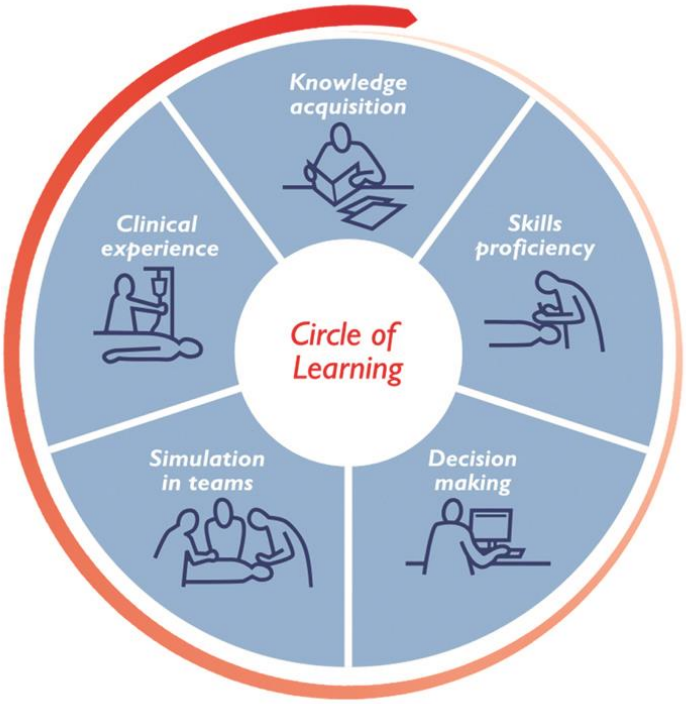
Kolb's model and theory are based on Dewey's "learning by doing theory". According to Dewey, the learning process is a continuous process, an improvement of knowledge where "learning by doing" is linking knowledge, skills and experience together. In both Kolb's and Dewey's theories, reflection is an important component in the learning process.

Another important pedagogical theory in medical simulation is *Blooms taxonomy* (46). The theory explains cognitive levels of learning. From lower levels where we remember facts through higher levels where we analyze, evaluate and create.

These theories have been implemented in a pedagogical model developed by Laerdal Medical AS called *The Circle of Learning* applied to medical simulation. The model was developed to enhance the understanding of simulation in a learning context, emphasizing that simulation is one of several interventions that should be considered to increase the probability of clinical implementation of knowledge, skills and behaviors. *The Circle of Learning* is often used as a

model for designing simulation-based education and training, and can also be used as a tool to identify and analyze learning needs (47).

Figure 2. The Circle of Learning applied to medical simulation.



(47)

Dieckmann et al. (48) presented four core perspectives of pedagogy in SBT in healthcare. In the book *Essential Simulation in Clinical Education*, chapter 4, they explain many of the behavioral psychological processes described in 2.6 into a simulation context. The aim was to break down the voluminous number of learning theories into core principles with practical examples on how these core principles might explain and be used as guidelines for designing SBT in healthcare. The authors emphasize that the learning process is a process of four core perspectives divided into three dimensions: Task, person and context. The learning process is described as the "adaption of the interplay between the task, person and context with the aim to create, recognize and use learning opportunities for the learners". The four sections are:

1. Behaviorism - how to influence observable behavior. Relation to practice: "The need for clearly defined learning goals and highly standardized training programs" with high psychological realism, standardized instruction formats and instructor training.

2. Cognitivist learning theories – explaining the cognitive processes of learning.

    a) Cognitive learning theory: Relation to practice: “Facilitation of activation of prior knowledge in elaborations, discussions, or seeking the boundaries of applicability of new information” through structured debrief.

    b) Cognitive load theory, focusing on learning complex skills by training strategies efficient for reducing the complexity.

3. The humanistic perspective – emphasizes learning as a possibility for human growth, where discussions and observational learning play an important role in developing self-efficacy. Relation to practice: Discussions and observational learning are important in developing different models for problem solving, learning from others experience. Knowledge and skills about debriefing and facilitation of discussions are important in the training of instructors.
4. Social learning theories - combining behaviorism, cognitivist learning theories and humanistic perspective focusing on how social behavior in groups e.g. a workforce, affects motivation, discussion and observational learning. Social learning theories are also explained as an efficient method for modification of behavior from a behavioral psychological perspective as described in 2.6. Relation to practice: The enablers and constrictors of learning and behavioral modification are related to social interactions with other people. Knowledge about human - and system factors, how an instructor can address them and facilitate exploration of how these factors affected the scenario and the participants. This is also an important part of the education of facilitators.

2.4 Quality indicators for clinical studies

The evidence of the findings in clinical studies should be rated according to its quality. A widely adopted tool for grading the quality of evidence and for making recommendations is GRADE (Grading of Recommendations, Assessment, Development and Evaluations), a framework for developing and presenting summaries of evidence and a systematic approach for making clinical practice recommendations (49). The GRADE framework evaluates the risk of bias, imprecision, inconsistency, indirectness, and publication bias and divides the quality of evidence into four categories; High, moderate, low and very low (Fig 3).

Figure 3: Summary of GRADE

A summary of GRADE's approach to rating quality of evidence

Study design	Initial quality of a body of evidence	Lower if	Higher if	Quality of a body of evidence
Randomized trials	High ⟹	Risk of Bias −1 Serious −2 Very serious Inconsistency −1 Serious −2 Very serious Indirectness −1 Serious −2 Very serious Imprecision −1 Serious −2 Very serious Publication bias −1 Likely −2 Very likely	Large effect +1 Large +2 Very large Dose response +1 Evidence of a gradient All plausible residual confounding +1 Would reduce a demonstrated effect +1 Would suggest a spurious effect if no effect was observed	High (four plus: ⊕⊕⊕⊕) Moderate (three plus: ⊕⊕⊕◯) Low (two plus: ⊕⊕◯◯) Very low (one plus: ⊕◯◯◯)
Observational studies	Low ⟹			

(50)

Bias can be defined as "a tendency which prevents unprejudiced consideration of a question" (51) Bias can be caused by several methodological factors like: lack of randomization, allocation concealment or blinding. Overestimation, exaggerated or missing harmful effects could be the results due to: a) intention to treat analyses are not performed b) reporting of interim results from trials that are cut short c) large attrition during follow up (participants drops out of the study population in follow up evaluation after intervention).

Rating according to the GRADE framework would be beyond the scope for this review. According to the aim, this study would rather focus on the quality of *documenting* the intervention in the included studies than the intervention itself.

Since there are no recognized or validated tools for grading the quality in simulation research, a grading system based on the reporting guidelines for healthcare simulation research by Cheng et al. (30) was designed for this thesis, based on four of the defined quality indicators according to Cheng et al.:

- Description of scenario and learning objectives.
- Instructional design; duration and frequency/repetitions.
- Feedback/debriefing; structure/method, duration, facilitator characteristics.
- Time used for intervention, repetitions, education of facilitators etc.

The four indicators were then transformed into seven identifiable indicators applied for assessing the quality of all included studies. Each indicator was weighted equally, counted as 1 of 7 possible, resulting in a score from 0 to 7.

1. Description of scenario
2. Description of learning objectives
3. Description of simulation structure/method
4. Duration of implementation
5. Description of facilitator simulation education and experience
6. Description of time used for each simulation intervention
7. Repeated simulation training

2.5 Human errors and adverse events

As mentioned in 1.1, the article *To err is human* (1) was published by the US Institute of medicine in 1999, revealing that preventable adverse events were the 8th leading cause of deaths in the United States. Between 44 000 and 98 000 deaths in U.S. hospitals, representing 2-4 % of the overall hospital admissions, could be linked to preventable medical errors. The report defines adverse events as "an injury caused by medical management rather than by the underlying disease or condition of the patient" They pointed to human errors as cause in a "sizable proportion of adverse events". The article also stated that errors most often occur due to multiple contributing factors, including errors to due human factors, often referred to as non-technical skills.

The consequences of a medical error or an adverse event might be serious for the patient involved. As discussed 1.3, the reports from WHO states (3, 4) the risk of being hurt or killed from adverse events and medical errors are quite high compared to other high risk industries. In a global perspective, as many as 4 of 10 patients are harmed in primary or outpatient care. 134 million adverse events occur annually in low- or middle-income countries, estimated to contribute to 2,6 million annual deaths. Medical errors are the third leading cause of death in the United States, and in United Kingdom it is estimated that one harmed patient is reported every 35 seconds. These facts led to the WHO declaration of the safety of health care to be a "major global concern." (3)

Even if healthcare systems differ a lot from country to country, the causes of harm and often the solutions as well, are quite similar. The WHO (10) patient safety curriculum guide

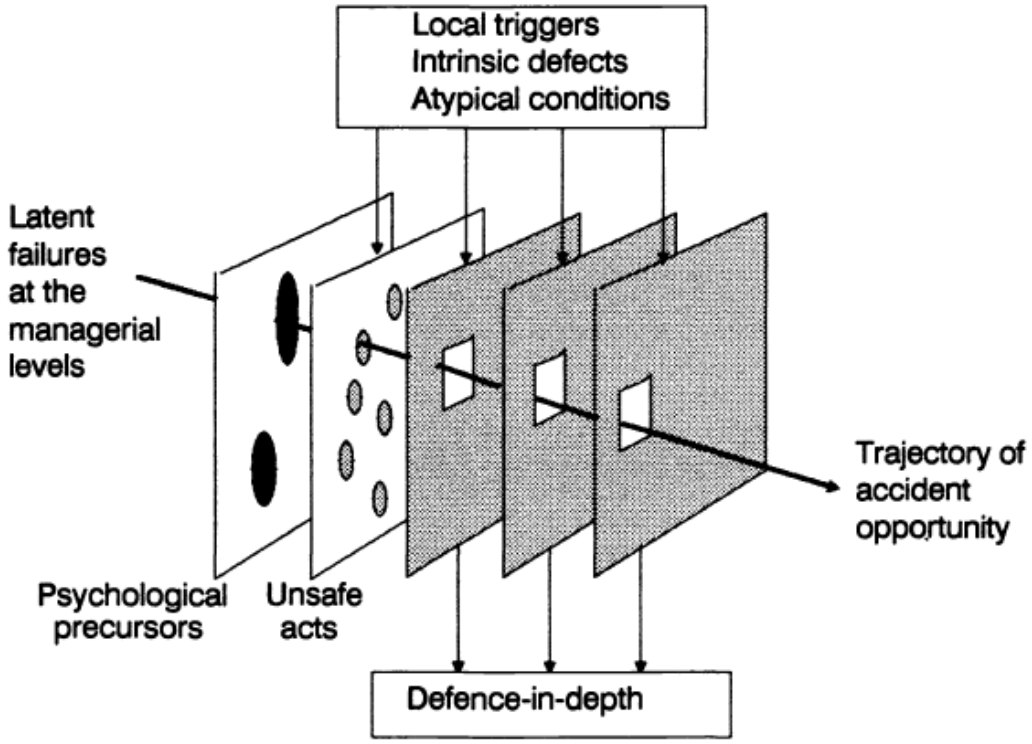
highlights the importance of applying human factors in the HCP educational curriculums to design and improve systems to reduce number of adverse events.

Human errors in health care - medical errors - are described as complex and multifactorial by nature. James Reason (52) stated in his book *Human errors* in 1990 that "errors depend on two kinds of failures; either the correct action does not proceed as intended (an error of execution) or the original intended action is not correct (an error of planning)". In the same book, Reason also presented a model explaining the dynamics of accident causation (Fig. 4). The model have later been modified and are often called "The swiss cheese model" (52). In this model, the last of several barriers is the healthcare provider (HCP). To make this last barrier against medical errors as strong as possible, HCP working as individuals or within medical teams need to understand the complexity of medical errors and work together in a way that reduces the risk of errors as opposed to increasing it.

Vincent et.al (8) introduced a theory of seven levels of safety to explain this complexity, and they pointed out pitches to be focused on in order to create, train and test defenses and barriers. Vincent et al. described both individual and team factors in these seven levels. Flin et.al (35) have described the characteristics of individual and team factors as being task management, teamwork, leadership, situational awareness and decision making. These are defined as nontechnical skills, and are considered as critical factors for reducing risk of errors (35).

SBT has been introduced as an efficient way of training nontechnical skills and teamwork in a safe environment. With the proper safety measures at place, there should be no risk for patient or provider (2, 53). However, this way of training is resource intensive, and might be questioned due to economic priorities and availability of resources.

Figure 4. The dynamics of accident causation:



(52)

For the EMS a systematic review (54) investigated and identified seven themes as main threats to patient safety in the EMS: 1.clinical judgement 2.adverse events and error reporting, 3.communications, 4.land vehicle safety, 5.aircraft safety, 6.interfacility transport and 7.intubation.

2.6 Culture and safety

Safety is closely connected to behavior (52, 55). Safety attitudes and behaviors are also key components in teamwork. Hollnagel et al. (56) defined safety in the book *Resilient health care* to be "a condition where nothing goes wrong" and follows up with an more realistic definition since it is impossible to create systems where nothing goes wrong "a condition where the number of things that go wrong is acceptably small." For HCP it is hard to accept that they make mistakes, but this is necessary to be able to design systems and build cultures that can enable health care to reduce the number of adverse events, injuries and deaths due to medical errors.

A safety culture has several layers and are more than the organizational focus to or willingness to have good systems for safety. As Reason (52) stated in his book *Human Errors,* the HCP is the last safety barrier in prevention of errors. Hence the HCP attitudes and behavior will be of importance in prevention of medical errors. Hudson (57) discussed and compared safety cultures and systems from other high risk industry to health care. Some relevant differences between health care and other high-risk industries are:

a) The consequences of an accident in health care: one patient injured or killed. In aviation or nuclear industry, the numbers of potential casualties are much higher.
b) Consequences for the individuals working in these industries, an aviation accident puts the flight crew themselves at risk of being injured or killed in an accident. For HCP an incident has no direct consequences for the responsible HCP. The risk of harm or death is on the patient.
c) From an organizational point of view, an aviation or a nuclear accident will have bigger consequences for the organization compared to a hospital organization after a single patient accident.

Hollnagel et al. (58) presented a view on safety and the causes of adverse events in the introduction of the safety II concept, where the idea is to seek to understand why things usually go right instead of the traditional safety approaches that focus on finding the causes for adverse outcomes.

Despite the fact stated by the WHO (4) that adverse events are one of the ten leading causes of death, most health care interventions are delivered and performed successfully. The adverse events are the exceptions. In their recent book, *Resilient Health Care* Hollnagel et al. (56) points to the fact that humans have a psychological tendency to "stop paying attention to something as soon as we get used to doing it." The practical implications of this, is that we have a hard time explaining how we perform successful routine procedures "every day." If we don't pay attention to the things that goes right, Hollnagel et al. state that an unavoidable consequence of this is an association of safety to the things that goes wrong. In psychological terms this is called habituation, and are closely related to human behavior. Loss of awareness have been linked to medical errors and adverse events, and are mentioned as a possible result from stress and reduced cognitive capacity (35) thus being highly relevant in medical emergency situations.

How people behave at work is regulated by factors like social norms, workplace regulations, ethical considerations, the relationship to colleagues, and management. All these factors can be defined as workplace culture. Bandura presented a social learning theory where he argues that behavior (Fig 5) partly is learned by the environment through the process of observational learning and mediational processes.

Figure 5. Behavior according to Bandura

He also explains how behavior might be modified. This kind of behavioral adaption is also occurring in our professional workplace. (59, 60).

Behavior modeling training suggests a model for practical training involving the following five steps:

1. Description of a set of behaviors to be learned for the participants
2. Modeling how these behaviors can be used effectively
3. Providing opportunity for practice the behaviors, rehearsal and repetition. The objective is to facilitate cognitive organization and retention.
4. Providing feedback and social reinforcement after practice
5. Helping participants to transfer the behaviors to their practice/job.

Murthy et al. (41) did a study comparing SBT vs. traditional role play for non-medical call center trainees. This study is interesting and relevant in this context because of the evaluation of effectiveness of SBT as a behavioral modelling modality compared to traditional role-play training. The Murthy et al. study found simulation to outperform traditional role-play in several areas. Their explanation for this was that simulation was found to be an efficient tool for behavior modelling. Simulation training improved the capacity to handle complex tasks by reducing the cognitive load and, the realistic context made transfer to their real-life job easier. They highlight three important environmental factors favoring simulation as a training modality;

- Realistic context where efforts are made to make a realistic presentation of the patient, where the equipment is what you would normally use and where you actually do the

interventions you normally would do following all steps in procedures etc. This adds realism to the perception of time to the context in general.

- Guaranteed feedback in a structured way. The training of simulation facilitators focuses on structure in debriefing, how different questioning techniques can be used to stimulate the participants' reflection and how direct feedback can be provided.
- Paced learning gives the participant the opportunity to continuously develop and refine knowledge, skills and performance. They can practice, backtrack and repeat as needed. This allows for a more rapid progression from novice to experts.

The importance of feedback as a behavioral modelling tool is also highlighted by Ericsson et al. (61) in the deliberate practice theoretical framework developed to develop novices to become experts. The work provided evidence on potential and limits of environmental adaptation and learning. Feedback seems to be a critical component in modelling of behavior as this is highlighted in all the theories presented in this chapter.

2.7 Human factors/non-technical skills

In 1988, John Sweller (62) described the cognitive limitations during problem solving and how this effects learning. Pattern recognition, cognitive mental models (structures), automation of processes (long term memory) are all strategies used to unload workload for short term memory.

As pointed out in 1.5.1, in a medical emergency, the HCP will be occupied by processing the complexity in the ongoing situation and by finding the right response on the problems presented. Due to the experienced stress the cognitive capacity and the access to short-term memory is limited and general awareness may be reduced. This explains a phenomenon known as "tunnel vision", a state where details in the environment "disappear" and the HCP focuses only on what he/she perceives as the most important factor in the situation.

To understand contributing and cofounding factors to accidents, analyzes were conducted in different industrial sectors. Finding that up to 80 % of all investigated accidents were caused by human factors. These human factors were categorized into seven topics, defined as skills that contribute to reduce risk in high-risk settings as the EMS.

1. Situational awareness,
2. Decision-making,

3. Communication,
4. Teamwork
5. Leadership
6. Stress management
7. Coping with fatigue.

These skills were explained in relation to their potential to contribute to accidents, and how the skills could be trained and used to reduce the risk of errors (35). Simulation is identified as a suitable tool to train these skills and improve patient safety (2).

There has been leveled criticism against the theories and work of Flin et al. arguing that the CRM approach to human factors as a tool for reducing risk of human errors is an over-simplifying of the topic and that the causes of human errors are way more complex than the seven topics identified by Flin et al.(35). Sidney Dekker (63) even doubted the term *human error* to exist other than as a "convenient but misleading explanatory construct", and he is not alone. His book *The field guide to understanding human error* brings forward the ideas from other prominent critics of existing human error terms and theories like David Woods, Erik Hollnagel, Nancy Leveson, John Flach, Richard Cook and Jens Rasmussen. They all represent a new view on human errors, where human error is seen as "a symptom of deeper trouble." (Fig 6).

Figure 6. Old view vs. new view of human errors according to Dekker:

Old View	New View
'Human error' is the *cause* of trouble	What we call 'human error' is a *symptom* of deeper trouble
'Human error' is a separate category of behavior, to be feared and fought	'Human error' is an attribution, a judgment that *we* make after the fact
'Human error' is the target; people's behavior is the problem we need to control	Behavior is systematically connected to features of people's tools, tasks and operating environment
'Human error' is something to declare war on. People need to practice perfection	'Human error' is information about how people have learned to cope (successfully or not) with complexities and contradictions of real work
'Human error' is a simple problem. Once all systems are in place, just get people to pay attention and comply	A 'human error' problem is at least as complex as the organization that helps create it
With tighter procedures, compliance, technology and supervision, we can reduce the 'human error' problem	With better understanding of the messy details of people's daily work, we can find ways to make it better for them
We can, and must, achieve zero errors, zero injuries, zero accidents	We can, and must, enhance the resilience of our people and organization

(63)

2.8 Teams and teamwork

Salas et al., (64) defined teams as being "social entities composed of members with high task independency, shared and valued common goals. They are usually organized hierarchically." Baker et al., (65) *Medical Teamwork and Patient Safety: The Evidence-Based Relation*, defined teams to have the following five characteristics:

1. Consists of two people or more
2. Team members are assigned roles, tasks and works towards a common goal
3. Team makes decisions
4. Contains specialized knowledge and skills, performs frequently under high workload.
5. Teams need coordination due to high degree of task interdependency.

The authors aimed to identify core competencies needed to enhance team efficiency and stated that "simply installing a team structure in an organization does not automatically result in effective teamwork. Effective team performance requires team members' willingness to

cooperate for a shared goal." The core competencies presented in this article are divided into 12 knowledge and 7 attitude competencies.

According to Salas (64), teams are characterized by the need for coordination and cooperation since task demands change during the process. Salas also used findings from previous team research to create an evidence-based framework for team skills and identified five core team competencies:

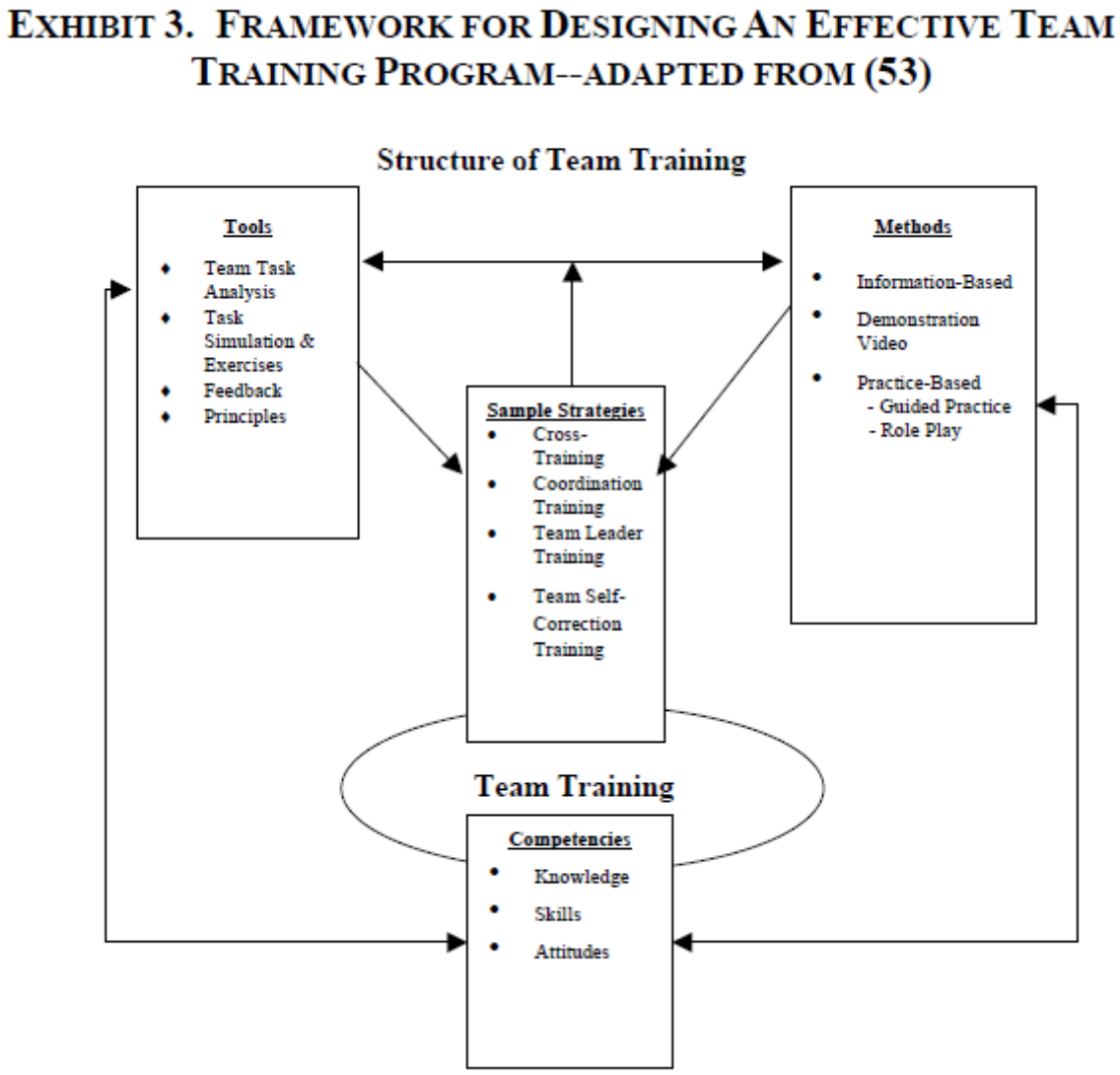
1. Team leadership
2. Mutual performance monitoring
3. Backup behavior
4. Adaptability
5. Team orientation

The attitude competencies were not a part of the five core competencies identified by Salas. Baker et al. defined team attitude competencies as "internal states that influence a team member's choices or decisions to act in a particular way" They also claimed that these attitudes might have a significant effect on the teamwork. Baker et a. identified the following attitude competencies:

1. Team orientation (morale)
2. Collective efficiency
3. Shared vision
4. Team cohesion
5. Mutual trust
6. Collective orientation
7. Importance of teamwork

SBT is recommended as one of several training strategies for training of these core competencies in teamwork and particulary for teams who will perform under stress, e.g. emergency medical teams. Baker et al. presented a framework (Fig 7) for designing an effective training program. This model is relevant for this review and complements the pedagogical models presented in 2.3. As the Circle of learning (47) the model clearly emphasizes that training needs a holistic approach, to optimize learning potential. A major concern for Baker et al. is the lack of interdisciplinary team training both in clinical practice and in educational programs despite a recognized need to coordinate patient care to improve patient safety.

Figure 7 Framework for team training according to Baker et al.



(65)

Organizations are increasingly using teams as a work group structure to solve complex tasks (66). J. Richard Hackman was one of the pioneers of team theories and research. In his 1998 article *Why Teams Don't Work* (67) he also described pitfalls for organizations when it comes to teamwork as for example mistake no. 6: "Assume that members already have all the skills they need to work well as a team." Hackman states that not many health care educational programs include team skills as an integrated part of the training, resulting in team members that lack the core competencies in team work mentioned above. Linking this information to the information in the previous section about human factors and cofounding factors to accidents and errors, the lack of core competencies in teamwork could represent a patient safety issue.

As stated before, Naik & Brien (2) found simulation to be an efficient way of training practical team skills. They also pointed to a possible explanation for the lack of focus on teamwork and non-technical skills in medical educational programs: There is no easy access to written curriculums, and there is a lack of competency in this area amongst most teachers and "you cannot teach what you do not know."

2.9 The Kirkpatrick framework

The Kirkpatrick Model (Fig 8) is probably the best known and validated model for analyzing and evaluating the results of training and educational programs (7, 68, 69):

Figure 8. Kirkpatrick's model of evaluation of training and education.

1. Reaction

   How did the participants perceive the training? Common form of assessment is quantifiable anonymous evaluation forms.
2. Learning

   Favorable reactions and enthusiasm does not equal learning and the amount of learning. To document learning, it needs to be measured in a quantifiable way. Before and after measurements are often used to measure the learning effect from the training intervention. Results are often analyzed in a statistical model to prove learning in terms of correlation and level of confidence.
3. Change in behavior

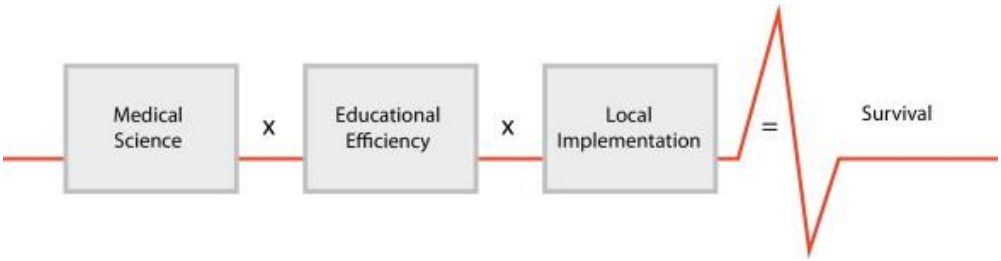
   Even if learning has been proved, it does not necessarily change the way people behave in situations where the training should be implemented. Kirkpatrick recommends appraisal of performance of one or more observers.
4. Change in results as measured for example in patient outcome or numbers of adverse events before and after.

2.10 The formula for survival

The theoretical connection between medical efficiency and patient outcome was published in the article "The formula for survival in resuscitation" (70). The formula (Fig 9) was first

presented in 2003 by the International Liaison Committee on Resuscitation (ILCOR) Advisory Statement on Education and Resuscitation as a hypothetical formula:

Figure 9. The formula for survival.



(70)

Three factors where highlighted as being interactive factors forming multiplicands determining survival from resuscitation:

1. Medical science: The quality of evidence behind guidelines.
2. Educational efficiency: Efficient education of the patient caregivers.
3. Local implementation: A well-functioning local chain of survival.

The hypothesis was tested in EMS systems over the years, and in 2013 Soreide et al. published the article presenting evidence of higher survival rates in systems using the formula as a system governance. From the results, the authors produced a theoretical example to show the potential of extra lives saved (Fig 10):

Figure 10. Example of practical use of the Formula for survival:

	Medical science	Educational efficiency	Local implementation rate (fraction of patients)	Number of EXTRA survivors per 100 treated
Best case	Constant A	Constant B	1.0	15
Real world	Constant A	Constant B	0.8	12
Beginners	Constant A	Constant B	0.4	6
Worst case	Constant A	Constant B	0.0	0

In this theoretical example the medical science (and thereby treatment recommendations) and educational efficiency are kept constant. The survival with no implementation of therapeutic hypothermia is set at 30% and with 100% implementation at 45%.23, 52

(70)

The model illustrates the importance of research on educational efficiency and is part of the background for the research question for this review article.

3.0 Methodical approach

The following chapter describes ambitions, the scientific methodology and design of this study. Rationale for the choices made are also explained.

The ambition was to answer the research question using scientific methods that a) meets publication standards b) includes recent and relevant published studies c) discusses the causes for conflicting results d) searches for evidence on how quality of SBT and different SBT designs influences the efficiency of the SBT intervention.

Magdalene Thomassen (71) states “Science can be described as a systematical exploration of reality” she also defines theory in a scientific philosophic perspective to be: “An abstract simplification of the reality. Describing and reasoning regular relationships between phenomenon’s, and structuring the facts into a meaningful holistic overview. Theories explain something general with the intention of explaining or increasing the understanding of a

phenomenon. “Scientific work has to meet certain criteria, most importantly it has to be verifiable, replicable and to make sense” (71).

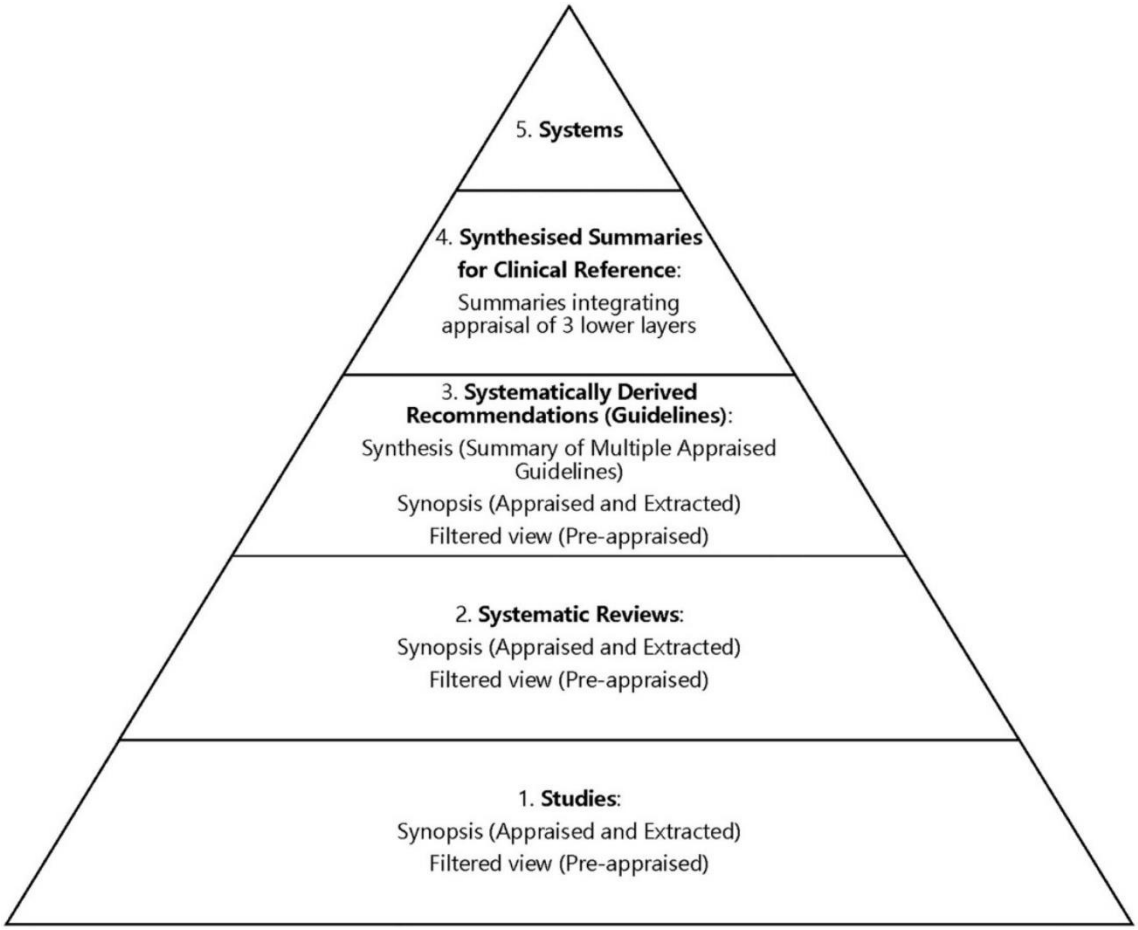
As medical care providers, we are encouraged to implement an evidence-based practice for the best of our patients. Sackett et al. (72) defined evidence based medicine: “It’s about integrating individual clinical expertise and the best external evidence”.

3.1 Research methods and design

The objective of this thesis was to search for answers to the research question whether SBT for health care emergency response teams leads to changes in patient outcome. Further analyses to investigate whether the training might have affected the reported mortality rates because of quality improvement, or other causes was not being differentiated or investigated in this study. The scientific method best suited to investigate the research question asked, was found to be; a systematic literature review.

In the process of initiating this study, it was important to get an overview of existing research or evidence from SBT. The evidence pyramid was used to map the level of evidence. Since having already identified and dismissed existing systematic reviews for reasons explained in 1.4, this study searched for evidence at lower levels in the pyramid, level 1- single studies with the aim of summarize the findings and report them as level 2 evidence.

Figure 11 The EBHC evidence pyramid



(73)

3.1.1 Why perform systematic literature reviews?

A systematic literature review is a form of evidence synthesis, an attempt to integrate empirical data and produce statements to guide decision making, identify knowledge gaps. By stating the review processes, element of arbitrariness is reduced, enabling other reviewers to replicate the review and its conclusions. Characteristics of a systematic review include a study protocol, a formal research question, eligibility criteria for the inclusion of studies, methodological systematic searches for studies, screening of publications against a priori criteria, formalized appraisals, assessment of scientific quality and risk of bias and explicit methods to combine findings (74). The PRISMA guidelines (Preferred Reporting Items for

Systematic Reviews and Meta-Analyses) are an established framework for systematic literature reviews, stating that systematic literature reviews are important in healthcare for clinicians to keep up to date and for granting credit to institutions and agencies to justify further research.
Scientific methodology, hierarchy and rationale for different methodologies (75).

The PRISMA guidelines were chosen as methodical framework for this study. PRISMA is a reporting and methodological guideline developed to assist in the reporting and performing systematic literature reviews in general. It is a widely accepted and used as standard for reporting and conducting systematic reviews (76) while other guidelines are tailored to specific types of systematic reviews and were not found suitable here (77-79).

3.2 Protocol and registration

The PRISMA-P 2015 checklist was used to design the systematic review protocol (75, 80). The protocol was submitted and registered at the University of Stavanger, Faculty of health sciences.

3.3 Eligibility criteria

3.3.1 PICO

To keep the search as relevant as possible it was important to limit the search only to include search words and studies relevant to the research question. The aim of this process was to identify studies reporting SBT interventions for individual HCP and teams of all fields in medicine and the effects on K4 level changes

The PICO Model is a format to define a question in order to help the searcher in finding clinically relevant evidence in the literature by posing four questions. The PICO methodology has been proven to be more effective due to greater sensitivity and specificity compared with other methods for systematic review literature and was therefore considered to be the best search tool for this study (81). For this literature search, the following PICO was designed:

1. Patient/ Problem or Population – medical emergencies, medical emergency situations, critically sick, critically injured, trauma patients, cardiac arrest, airway emergencies, post-partum bleeding, stillbirth, massive bleeding, life threatening medical situations, Emergency

medical team, Medical emergency team, HEMS, Rapid response teams (RRT), Patient care team, Air ambulance, Emergency helicopter, Ambulance, Emergency.

2. Intervention – Simulation based training

Characteristics of medical simulation is a clinical scenario followed by a structured debrief with the aim to reflect and learn from the simulated experience. The pedagogical framework for this model is Kolb's experimental learning theory and learning cycle (44).

3. Control/comparison/comparator – not relevant

4. Outcome –Improved mortality rates, reduction of medical errors, reduction in numbers of adverse events.

Efficient training could produce results in several ways; quality improvement which again could lead to increased quality of diagnosis or treatment as well as reduction of adverse events potentially affecting outcome and reduced mortality rates. This will not be investigated in this literature review and might be a limitation of the study.

In addition to the PICO criteria, eligible studies had to be published in English as full text original article. Abstracts, books, theses and conference proceedings were excluded, as were studies involving students.

3.4 Information sources

The search was executed with the help of a professional librarian and conducted in the following databases: MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials, PubMed, CINAHL, Epistemonikos and Google Scholar. Systematic reviews were excluded. Studies were also identified by scanning of reference lists of articles and systematic reviews. Searches were performed October 28-31, 2019

3.5 Search

The EMBASE and MEDLINE searches provided the most relevant studies and are the sources of most of the included studies from the literature search and are presented under. For the other databases the search protocol was adjusted according to requirements. Search strategies can be found in appendix 7.2.

Database: Embase <1974 to 2019 October 28> Search Strategy:

---

1 (emergency medical team* or medical emergency team* or HEMS or rapid response team* or patient care team* or air ambulance* or emergency helicopt* or airway response team* or medical emergenc* or medical emergency situation* or critically sick or critically injured or trauma patient* or cardiac arrest or airway emergenc* or post-partum bleeding or stillbirth* or massive bleeding or life threatening medical situation*).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (114333)

2 (((simulation adj3 training) or (patient adj3 simulation*)).mp. or simulation training/ or scenario based training.mp.) not (computer simulation*.mp. or computer simulation/) [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (11163)

3 mortality/ or mortality rate/ or (mortality or ((improved or reduc*) adj3 (mortality adj3 rat*))).mp. (1396012)

4 1 and 2 and 3 (66)

5 exp medical error/ or error/ or (error* or mistake*).mp. (700682)

6 4 and 5 (7)

7 (((systematic* or literature) adj3 (overview* or review* or search*)) or meta analys*).ti,ab. (657367)

8 6 and 7 (0)

9 4 and 7 (3)

Database: Embase <1974 to 2019 October 28> Search Strategy:

---

1 (emergency medical team* or medical emergency team* or HEMS or rapid response team* or patient care team* or air
ambulance* or emergency helicopt* or airway response team* or medical emergenc* or medical emergency situation* or critically sick or critically injured or trauma patient* or cardiac arrest or airway emergenc* or post-partum bleeding or stillbirth* or massive bleeding or life threatening medical situation*).mp. [mp=title, abstract, heading word, drug trade name,

original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (114333)

2 (((simulation adj3 training) or (patient adj3 simulation*)).mp. or simulation training/ or scenario based
training.mp.) not (computer simulation*.mp. or computer simulation/) [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (11163)

3 mortality/ or mortality rate/ or (mortality or ((improved or reduc*) adj3 (mortality adj3 rat*))).mp. (1396012)

4 1 and 2 and 3 (66)

5 exp medical error/ or error/ or (error* or mistake*).mp. (700682)

6 4 and 5 (7)

****

3.6 Study selection

3.6.1 Selection process

References was exported to EndNote X8 and scanned for duplicates which were removed. Systematic literature reviews were manually scanned for relevant articles, they were added to the selection process as additional records identified through other sources.

Three independent reviewers screened title, abstracts and full text articles. The reviewers have relevant clinical and simulation experience. Two are paramedics and one consultant in anaesthesia coming from Norway and Finland. The reviewers are all members of the international EuSim faculty. EuSim is an organisation where leading European simulation centers collaborate with a common goal to support the use of simulation in health care for improving patient safety and quality of care, hence the reviewers have broad insight in simulation theory and practice. One reviewer has a PhD, one has a master's degree and the third one is the author of this thesis. For the screening process, the Rayyan QRCI tool was used (82). This software was developed to help expedite the abstract screening process for systematic literature reviews. It is reported to be a useful tool (83). Three reviewers screened abstracts and full text articles against eligibility criterions independently. Disagreements were solved by consensus or majority decisions.

3.7 Data collection process

Information was extracted from each full text article and a data extraction form was generated, based on the information that was relevant to review the studies according to the research question and PICO for this review. This form can be found in the attachments as required by the PRISMA guidelines.

3.8 Quality assessment of included studies

Quality assessment of included studies was performed using the grading system presented under 2.4.

3.9 Summary measures

Changes on K4 level: To present summary measures as included in the PRISMA checklist, all studies were evaluated whether there were changes found in mortality rates, number of adverse events or medical errors. Changes were evaluated as reduction, increase or no changes.

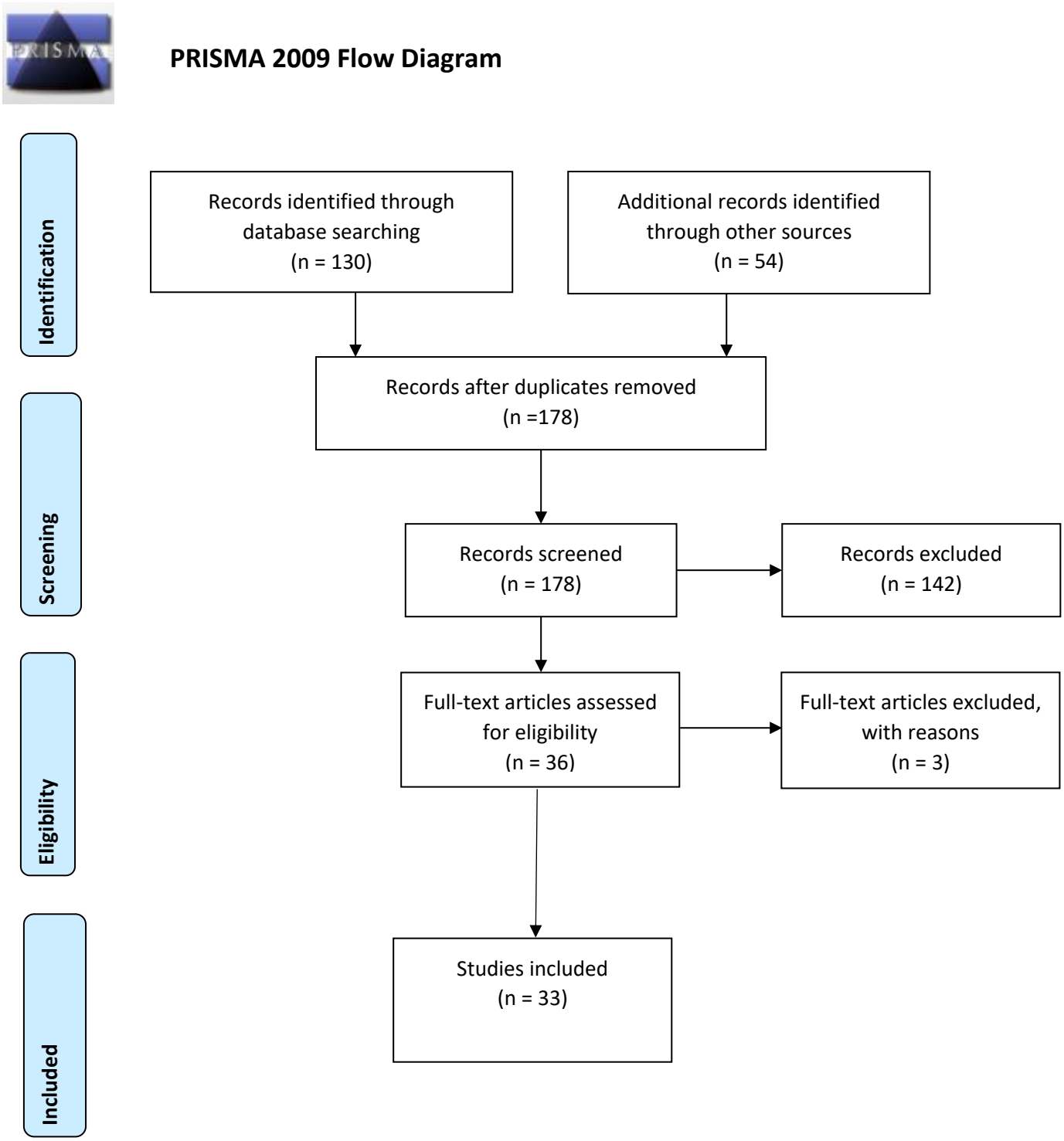
4.0 Results

In this chapter, the results from the study selection and data extraction will be presented. The results were analyzed with the aim to produce results that could give answer to the research question. This included quality assessment and quality rating of the simulation interventions, investigation of changes at K4 level for mortality rates, number of adverse events or number of medical errors. Sub analysis were performed to investigate if there were any SBT design similarities in studies with or without K4 level changes and whether study design affected sustainability of changes.

4.1 Study selection

The result of the study selection is presented in the Prisma flow diagram (Figure 12)

Figure 12. Prisma 2009 Flow Diagram



Of the studies identified for the screening process 54 (29%) were additional studies identified through the manual screening of articles or reviews. There were no duplicates identified. Three of the studies included through the abstract review process were excluded after the full text analysis due to lack of measurements of patient outcome (84-86).

4.2 Study characteristics

Of the included studies, 32 were observational cohort studies, prospective or retrospective, while one was a randomized controlled trial (RCT) (14). Four were longitudinal studies (16, 20, 87, 88). Two studies were longitudinal follow up studies of earlier interventions (87, 88). All studies were conducted in hospital settings, from both urban and rural areas and from low- and high resource health care systems. No pre-hospital/EMS studies were found.

4.3 Results of individual studies

A full data extraction table containing information about design, setting, population, intervention, outcome measure, K4 level results and notes for each of the included studies can be found in appendix 7.1. The focus for the results was, according to PICO, about the K4 level changes regarding patient outcome as measured by reduced mortality and reduced number of adverse events and medical errors.

Two studies were identified measuring at an even higher level, above the original Kirkpatrick levels, introducing K5 level measuring return of investment (85, 88). Both these studies were longitudinal follow ups of included studies originally measuring and proving K4 level effects. Theilen et al. was included in the review since the study followed up K4 level results from the initial study, while Van de Ven et al. was left out of the review, but included into the discussion.

Table 1 presents the quality assessment in reporting the SBT interventions as described under 2.2.

One positive finding according to the table is rated 1 point. Maximum score is 7/7 while minimum score is 0/7.

Table 1: Quality assessment in reporting the SBT interventions

Study ID	Description of Scenario	Description of Learning objectives	Description of simulation structure/method	Duration of educational intervention	Description of facilitator sim education and experience	Description of time used for each simulation intervention	Repetitions?	Rating
Ajmi et al. 2019	Yes, in supplemental material	Yes, in article	Yes, in supplemental material	February 2017 to present (still ongoing)	yes	60 minutes	Yes, weekly	**Rating 7/7**
Andreatta et al. 2011	No	Yes	Yes	4 years	No	No	Yes, at least monthly	**Rating 4/7**
Arabi et al. 2018	No	No	No (HBB concept)	6 months	Yes	No	No	**Rating 2/7**
Bellad et al. 2016	No	No	No (HBB concept)	No	Yes	No	Yes (low dose high frequency)	**Rating 2/7**
Braddock et al. 2015	No	Yes	Yes	1 year	Yes	30 minutes	Yes	**Rating 6/7**
Carlo et al. 2010	No	No	No	6 years	Yes	No	Yes	**Rating 3/7**
Chang et al. 2019	No	No	Yes	2 Years	No	No	No	**Rating 2/7**
Christensen et al. 2016	No	Yes	No	22 months	No	No	No	**Rating 2/7**
Copson et al. 2017	No	Yes	Yes	2006 – to present	No	No	No	**Rating 3/7**
Dillon et al. 2018	No	No	Yes	4 years	No	30 min.	Yes	**Rating 4/7**
Draycott et al. 2008	No	Yes	No	8 years	No	No	No	**Rating 2/7**
Fransen et al. 2017	Yes	Yes	Yes	14 months	Yes	50 min	NO	**Rating 6/7**
Fuhrman et al. 2009	No	Yes	Yes	5 months	Yes	90-105 min	No	**Rating 5/7**
Goudar et al. 2013	No	No	No	10 months	Yes	No	Yes	**Rating 3/7**
Inglis et al 2011	No	No	No	2 months	No	No	No	**Rating 1/7**
Jung et al. 2016	No	No	No	6 months	No	No	Yes	**Rating 2/7**
Knight et al. 2014	No	Yes	Yes	6 months	No	No	Yes	**Rating 4/6**
Mduma et al. 2015	No	Yes	Yes	12 months	Yes	Yes	Yes	**Rating 6/7**
Msemo et al. 2013	No	Yes	Yes	12 months	Yes	No	No	**Rating 4/7**
Mehta et al. 2013	Yes	No	Yes	24 months	No	60 min	No	**Rating 4/7**
Neily et al. 2010	No	Yes	Yes	24 months	No	No	Yes	**Rating 4/7**
Nelissen et al. 2017	No	Yes	Yes	1 month	Yes	Yes	No	**Rating 5/7**
Phipps et al. 2012	No	Yes	Yes	4 months	Yes	120 min.	No	**Rating 5/7**
Riley et al. 2011	No	No	Yes	6 months	No	150-165 min.	Yes	**Rating 4/7**
Shoushtarian et al. 2014	No	No	No	24 months	Yes	No	No	**Rating 2/7**
Siassakos et al. 2009	No	Yes	Yes	24 months	No	No	Yes	**Rating 4/7**
Sodhi et al. 2015	No	No	No	12 months	Yes	No	Yes	**Rating 3/7**
Theilen et al 2013	No	Yes	Yes	12 months	No	120 min.	Yes	**Rating 5/7**

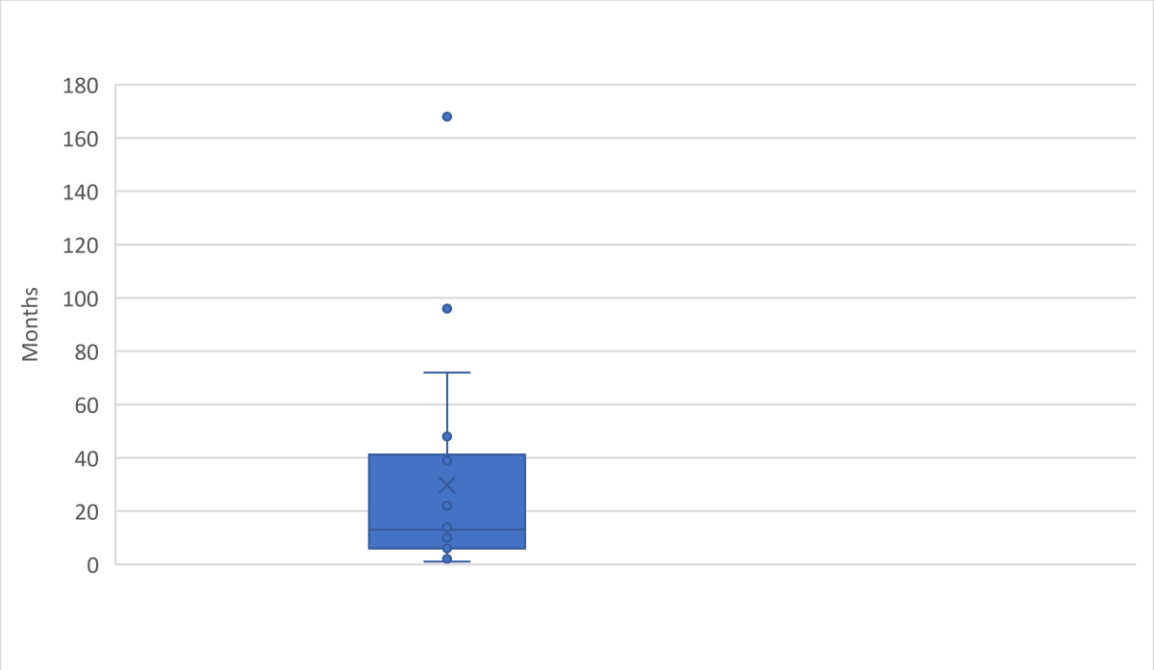
Theilen et al. 2017	No	Yes	Yes	24 months	No	120 min.	Yes	**Rating 5/7**
Van de Ven et al. 2017	Yes (in original article)	Yes	Yes	14 months	Yes	50 min	Yes	**Rating 7/7**
Wagner et al. 2012	No	No	No	24 months	No	No	Yes	**Rating 2/7**
Walker et al 2016	No	No	No	24 months (12 months follow up)	No	No	No	**Rating 1/7**
Wehbe, Janek 2013	Yes	Yes	Yes	6 months	Yes	No	No	**Rating 5/7**

In total, 19 studies (58%) reached a rating of 4/7 or more. The scenarios used in the SBT intervention were described in five studies (15%), while learning objectives were described in 19 studies (58%), and 21 (64%) studies described the simulation structure or method.

4.3.1 Duration of educational intervention

All but one study (97%) described the duration of the intervention, and median time was 14 months. The distribution of the duration of implementation periods for the included studies is presented in Fig.12, showing that there were two outliers with much longer durations of 98 respectively 170 months.

Figure 13. Duration of educational intervention



Facilitator experience was described in 16 studies (48%), 13 studies (39%) described the time used for each simulation and 18 studies (56%) repeated the SBT.

Table 2 presents the results of K4 level effects in relation to the grading scores:

Table 2 - K4 level results and quality rating

Quality rating	1/7	2/7	3/7	4/7	5/7	6/7	7/7
K4 level changes found	Inglis et al. 2011 Walker et al. 2016	Arabi et al. 2018 Chang et al. 2019 Draycott et al. 2008 Jung et al. 2016 Shoushtarian et al. 2014 Wagner et al. 2012 Bellad et al. 2016 Christensen et al. 2016	Carlo et al. 2010 Goudar et al. 2013 Sodhi et al. 2015	Andreatta et al. 2011 Dillon et al. 2018 Knight et al. 2014 Msemo et al. 2013 Neily et al. 2010 Riley et al. 2011 Siassakos et al. 2009 Mehta et al. 2013	Nelissen et al. 2017 Phipps et al. 2012 Theilen et al. 2013 Theilen et al. 2017	Braddock et al. 2015 Fransen et al. 2017 Mduma et al. 2015	Ajmi et al 2019 Van de Ven et al. 2017
No K4 level changes found			Copson et al. 2017		Fuhrman et al. 2009 Wehbe et al. 2013		

Of the included studies, 30 (91%) described significant K4 level improvements, while three studies (9%) found no K4 level changes. The results presented in Table 2 subsume mortality, adverse events and/or medical errors.

4.3.2 Changes in mortality rates

Of the included studies, 22 (67%) had mortality as one of the pre-defined outcomes; 18 of these 22 (82%) found reduced mortality rates, one did not document statistical evidence and four (18%) found no effect on mortality.

The findings for each study are presented in the list below:

1. Ajmi et al., 2019 (89) – reduced 90 days mortality from 9,1% to 3,5%.

2. Andreatta et al., 2011 (16) - increased survival rates from 33% to 50%.
3. Arabi et al., 2018 (90) – Reduced fresh stillbirth rates from 10.5 to 3,3 per 1000 births. Mortality reduction in newborns who received mouth to mouth ventilation before (50%) compared to newborns receiving bag-mask ventilation after intervention (11%). And a decrease in mortality rates for early neonatal deaths from 13,5 to 4,3 per 1000 live births.
4. Bellad et al., 2016 (91) – Improved survival. 46% reduction of stillbirths and 17% improvement of 7 days mortality of newborns weighing less than 2500g at one study cite, but no significant results at the other two centers in the study.
5. Braddock et al., 2015 (92) – The weighted risk from observed to expected mortality ratio decreased from 0,50 to 0,40.
6. Carlo et al., 2010 (93) – Significant reduction in the rate of stillbirth from 23/1000 to 15,9/1000 births.
7. Chang et al., 2019 (94) – Rate of maternal mortality from obstetric hemorrhage reduced significantly from 1,2% – 0,2%.
8. Christensen et al., 2016 (95) - Immediate post-code survival and survival from 71,1% to 67,3% did not translate into increased survival to discharge.
9. Dillon et al., 2018 (96) - Significantly more patients survived to discharge after RRT training was implemented.
10. Goudar et al., 2013 (97) – Reduced number of stillbirths from 3.0% – 2,3%. Fresh stillbirths decreased from 1,7% - 0,9%.
11. Jung et al., 2016 (98) – Unexpected mortality decreased from 21.9 to 17,4 per 1000 discharges. Overall mortality significantly decreased from 39,9 to 34,6 per 1000 discharges.
12. Knight et al., 2014 (99) – Post cardiac arrest survival to discharge increased from 40,3% to 60,9%.
13. Msemo et al., 2013 (20) – Reduction of 24h neonatal mortality by 47% and 24% reduction of fresh stillbirth.
14. Mduma et al., 2015 (100) - Reduction of 24 h mortality by 40%.
15. Mehta et al., 2013 (24) – Reported reduced trustwide impatient airway related mortality from three before intervention and none in a two-year period after intervention.

16. Neily et al., 2010 (15) - Observed reduction in 30 days mortality rate of 18% vs. 7% in control group. Risk adjusted annual mortality reduction was found to be 50% compared to control group.
17. Sodhi et al., 2015 (101) –Increased survival from 26,7% – 40,8% and increased survival to discharge ratio from 23,4% to 66,6%.
18. Walker et al., 2016 (14)– Hospital based neonatal mortality reduced by 40% found 8 months after intervention. No significant differences before/after training at 4 or 12 months.

4.3.3 Studies with no change in mortality rates

1. Copson et al., 2017 (102) - No significant changes in perinatal mortality.
2. Fransen et al., 2017 (103)– No changes in maternal and perinatal mortality.
3. Fuhrman et al., 2009 (104) - No changes in 30 and 180 days mortality for patients at risk.
4. Wehbe et al., 2013 (105) – No significant improvement measuring relationship of increased activation of RRT and changes in hospital mortality rates.

4.3.4 Changes in adverse events

Adverse events were measured in 17 (52%) of the included 33 studies and 15 of the 17 (88%) studies found a significant reduction in numbers, or in outcome score. Two of the 17 (11%) found no changes in numbers or outcome scores of adverse events.

1. Ajmi et al., 2019 (89) - Improvement in 90 days outcome post stroke with worst outcome reduced from 12,2% to 3,5%.
2. Braddock et al., 2015 (92) - Reduction of hospital acquired severe sepsis/septic shock from 1,78 to 0,64, decreased incidents of acute respiratory failure 2,44 to 0,43 (per 1000-unit discharges) and increased number of days between cases of severe sepsis/septic shock.
3. Draycott et al.,2008 (22) - Rate of obstetric brachial plexus injury decreased from 7% to 2,3%. Overall injuries reduced from 9,3% - 2,3%.
4. Fransen et al., 2017 (103) – Trauma due to shoulder dystocia reduced from 0,31% to 0,16%.

5. Inglis et al., 2011 (106) – Trauma due to shoulder dystocia reduced from 0,40 – 0,14 (per 1000?) and incidence of obstetric brachial plexus injury (OBPI) significantly reduced from 30% to 10,67%.
6. Phipps et al., 2012 (18) - Decrease in average adverse outcome index (AOI) score from 0,052 before to 0,043 after SBT intervention.
7. Nelissen et al., 2017 (107) – Reduction in incidence of post-partum hemorrhage from 2,1 before to 1,3 after training.
8. Riley et al., 2011 (21) – 37% reduction in perinatal morbidity.
9. Shoushtarian et al., 2014 (108) – Improvements in cord lactate levels.
10. Siassakos et al., 2009 (109) – Reduction in median time from diagnosis to delivery for umbilical cord prolapse from 25 to 14,5 minutes. And a significant increase in caesarean sections were recommended actions had been performed from 34,78% to 82,35%.
11. Theilen et al., 2013 (19) – Reducing time to recognize deteriorating patients from 4 hours to 1,5 hours, increasing number patients reviewed by a doctor from 45 to 76% and reduction of ICU admissions.
12. Theilen et al., 2017(88) – Sustained and even improved results after 3 years. 81% of patients being reviewed by a consultant, number of ICU admissions reduced from initially 56 to 32 after 3 years.
13. Van de Ven et al., 2017 (87) – Reduced trauma due to shoulder dystocia and improvement of invasive treatment for severe postpartum hemorrhage.
14. Wagner et al., 2012 (110) – Decrease in Adverse Outcome Index (MAOI) from 2% to 0,8%.
15. Walker et al., 2016 (14) - Lower number of cesarean deliveries in the simulation intervention group.

4.3.5 Studies with no changes in adverse events

1. Copson et al., 2017 (102) – No changes in intervals between diagnosis to delivery, incidence of documentation, incidence of ruptures.
2. Fuhrman et al., 2009 (104) – No changes in rate of nursing staff awareness of patients at risk.

4.3.6 Reduction of medical errors

No studies measured reduction or increase of medical errors.

4.3.7 Study compliance/adherence to recommended reporting guidelines after introduction in 2016

Reporting guidelines for SBT studies were implemented in 2016 (30). Sixteen studies were published after the implementation and rated regarding their adherence to these guidelines. Seven had a higher score than 3/7. Results are presented in Table 3.

Table 3 – Compliance to reporting guidelines of studies published after the implementation of reporting guidelines.

Quality rating	**1/7**	**2/7**	**3/7**		**4/7**		**5/7**	**6/7**	**7/7**
Study id.	Walker et al. 2016	Arabi et al. 2018 Bellad et al. 2016 Chang et al. 2019 Jung et al. 2016	Copson et al. 2017.		Dillon et al. 2018		Nielssen et al 2018 Theilen et al. 2013 Theilen et al. 2017	Fransen et al. 2017	Ajmi et al. 2019 Van de Ven et al. 2017

4.3.8 Repeated SBT and K4 level outcome

Of the 33 studies, 18 (55%) included repeated training sessions in their intervention while 15 (46%) did not as presented in table 1 column 7. All 18 studies found a significant K4 level effect.

4.3.9 Sustainable changes in outcome

Two studies found the longitudinal effect on patient outcome declining 3 months after the training intervention (87, 96), while three studies found the longitudinal effect to be maintained with repeated training (16, 20, 88)

5.0 Discussion

The discussion is presented in two chapters. 5.1 explains how the investigation to answer the research question were performed, how different lay out of the included studies might affect the results and the quality of evidence from this study, how studies were selected, and data extracted. The results are discussed in 5.2.

5.1 Description of the process

The objective of this thesis was to examine the research question: Can SBT for health care emergency response teams reduce mortality rates, number of adverse events and risk of errors? While the primary aim of this study was to summarize recently published research in an attempt to answer the research question in addition, an evaluation of the quality in reporting of the applied SBT interventions was performed, aiming to explore if there were any SBT design similarities in studies related to reported results and whether study design affected sustainability of changes.

The topic was chosen due to personal experience and lessons learned from both clinical work and SBT, that medical emergency teams should train to perform at a sufficient standard, and that an appropriate way to perform training may be via SBT. Chapter 2 outlined the theoretical background for the need of training, how simulation is used as a training method, and how the impact of an intervention can be measured by applying outcome levels as defined by Kirkpatrick (7, 69). It also described systems for rating the quality of research like the GRADE system. The main focus in this review was not on quality of evidence in the included studies but the quality in reporting what has been done. As no suitable or recognized tool were identified, a novel rating tool was developed based on the reporting guidelines by Cheng et al (30).

In Chapter 3 the applied methods are described. In the process of compiling an appropriate search algorithm most of the included studies were found to be observational cohort studies. The ideal scientific method to examine the effect of an intervention would be a RCT. However, Sackett et al. (72) wrote in their article *Evidence based medicine: what it is and what it isn't* that "evidence based medicine is not restricted to RCTs and meta-analysis, it involves tracking down the best external evidence with which to answer our clinical question." David Gaba (111) stated: "No industry in which human lives depend on the skilled

performance of responsible operators has waited for unequivocal proof of the benefits of simulation before embracing it." These statements are relevant and important when discussing studies and their contribution to knowledge.

Evidence from observational studies are lesser valued than evidence presented in RCTs. The effect of unaccounted confounders maybe higher in observational studies than in RCTs. Several studies have contrasted the results from RCTs and observational studies without finding systematical difference, refuting the assumption of observational studies producing untrustworthy results (112). Observational studies might produce reliable results.

Systematic literature reviews based on cohort studies will probably be the highest quality of evidence possible to generate for training-/simulation interventions. Based on this rationale and experience, research aiming to investigate the effect of training/simulation interventions will most likely be based on observational studies also for future review articles.

There are numerous publications about the effect of SBT on Kirkpatrick levels 1 to 3, but systematic literature reviews covering K4 level have been either performed some time ago and thus leave out recent and relevant literature or cover very narrow fields of medicine, as discussed in 1.4. To the best of my knowledge there are no recent reviews covering K4 level effects of SBT for medical emergency teams across different medical fields. Thus, in order to contribute new knowledge to this field, the search was designed to include studies reporting effects on at least K4 level.

5.1.1 Extraction and selection of studies included

The review process (described in 3.6) did not identify any systematic literature reviews investigating the impact of SBT on the K4 level for medical emergency teams in general. All included literature from the search performed for this study were original studies, mostly observational studies. When writing the PICO, K4 level effects needed to be identified and defined. The definitions were discussed between the three reviewers. Definitions needed to be precise, realistic and relevant. We decided it to be for this study mortality, adverse events and medical errors, as they reflect on patient survival and patient safety.

The search words were defined from the PICO. Knowledge and experience from the literature search conducted for the scoping review, clinical practice and medical simulation were important in this process. A librarian assisted in this process, using experience from similar

searches and results from the performed searches to adjust the search words. Our combined knowledge and experience were essential in the process of defining the search words used for identification of all relevant studies related to emergency medical teams and SBT reporting on K4 level results. The same procedure had to be repeated to every term described by the PICO. Each of these terms had to be supplemented by synonyms, alternative definitions, concrete examples etc. E.g. the term medical emergency had to be specified; critically sick, critically injured, trauma patient, cardiac arrest, airway emergency, post-partum bleeding, stillbirth, massive bleeding, life threatening medical situation. In this critically important process, imprecise, wrong definitions or missing words could result in loss of evidence.

The results from the performed literature search were presented and discussed with the two co-reviewers. Screening of abstracts and full-text studies process was performed as described in chapter 3.6. The results from the performed literature search were presented and discussed with the two co-reviewers.

Screening of abstracts and full-text studies process was performed as described in chapter 3.6. The Rayyan QRCI tool (82) was used in this process. None of the reviewers had any prior experience with the free web or mobile based tool. The user experience in the screening process was positive. The application was user friendly and intuitive to learn. Rayyan QRCI made it possible to easily share included references, abstracts and full text articles to all the reviewers. The Rayyan QRCI tool also provides the possibility of blinding the reviewers, so they don't know if the other reviewers have included or excluded studies. Monitoring of the progress was easy, when all review items had been evaluated. The results of the screening process where intuitively presented by the tool. The tool had been found useful by other researchers (83) and was also perceived to be very useful in the process for this review. One benefit was the elimination of potential inconvenience due to large geographical distances between the reviewers. Since the Rayyan QRCI tool is freeware, this makes the program available to all who might have the need for such tool. Of the included studies, 54 (29%) were not identified through the literature search but as described in 3.6.1 by manual screening of other systematic literature reviews. All studies were included in the Rayyan QRCI screening process regardless of how they were identified

Important assets to prevent loss of potentially important studies where the algorithm for screening described in 3.6 and the combined knowledge and experience of the 3 reviewers and the librarian, from scientific research, clinical and simulation practice.

Most of the 142 excluded studies were rejected because they focused on improvement, implementation of new protocols or equipment, and fewer focused on patient outcome.

After the inclusion process, a data extraction table were designed and relevant data from each study were included in this table. The data extraction table can be found in appendix 7.1. Relevant data was discussed amongst the reviewers, and was decided to be for this study: Study ID (author/year), study design, translated study design, setting, population, intervention, outcome measure and K4 level results. The extracted data were the source for the narrative analyses performed and presented in chapter 4.

While assessing the theoretical frameworks for this study, the GRADE system for rating quality of study design and performance was found to be beyond the scope of this review. As the included studies had all been peer-reviewed as part of the individual publication process for each study, performing an alternative quality review of the SBT intervention seemed more relevant to enable this study to answer the research question asked. Publications in clinical research should contain a description of the methods applied, allowing other researchers to repeat the study and confirm or reject the results. The SBT intervention should be a part of the method description. To evaluate the quality in reporting the SBT interventions a rating tool was developed based on the reporting guidelines from Cheng et al. (30) as described in 2.4 The objective was to increase insights in how the SBT design adheres to the principles presented in the theories explained in 2.2, 2.3 and 2.6. This kind of detailed information would again make it easier to see whether the SBT was designed in a way that would facilitate the full learning potential of SBT in line with the theories presented by Bandura et al., Dieckmann et al. Kolb et al. and Bloom (44, 46, 48, 59). Demonstrated by Msemo et al. (20) in their study where they found SBT to be an effective way of modeling clinical behavior.

5.2 Discussion of results

This review over recent literature about SBT included 33 studies. Of these were 32 observational studies and one was a RCT. When evaluating the quality in reporting of the intervention, 19 (58%) of the studies where rated 4/7 points or higher. Thirty studies (91%) described significant K4 level improvements, while three (9%) found no K4 level changes. All 18 studies (55%) that reported that the intervention included repeated SBT found positive results on the K4 level.

5.2.1 Study/research contexts

When searching for research in emergency medicine, it would be expected to find studies performed in typical arenas where these incidents happen and normally would be treated, namely in the pre-hospital environment. However, none of the studies included pre-hospital emergency teams. A possible explanation for this might be that pre-hospital medical emergencies are initially treated by EMS, but during the clinical cause following the initial handling of the medical emergency are treated in different wards resulting in a large number of confounding factors e.g. changes in protocols, new interventions or medications etc. Mduma et al.(100) found that K3 level changes, observed clinical behavior also led to K4 level changes, 40% reduction in mortality rates. There is a plausible relationship between K3 and K4 level changes, but how it correlates remains unclear.

Proven educational efficiency might be expected to have impact on patient outcome according to The formula for survival (70). According to the formula, studies proving educational efficiency, might have impact on survival. Proving educational efficiency through e.g. observations of clinical behavioral at K3 level, might implicate impact on higher Kirkpatrick levels. This is supported by the findings from the Mduma et al. (100) study. The design of this study might be a model that could be useful to apply for future pre-hospital studies for evaluation of effects on patient outcome after SBT.

5.2.3 K4 effects

The analysis of the results gave an indication of K4 level changes to better outcomes after SBT. Ninety percent of the included studies found significant K4 level improvements. For the studies investigating the impact of SBT on mortality rates, 82% found positive results. While 88% of the studies investigating adverse events found reduction in numbers or improved adverse event scores. Even if the number of studies identified and included for this systematic review were low, the share of reduced mortality rates and adverse events on the K4 level are high within the included materials. These findings could support the expected and claimed efficiency of simulation by Gaba, Naik and Brien and the recommendation given in the *To err is human* report (1, 2, 6). The importance of educational efficiency and the potential effects on increased survival were presented and discussed in 2.10 The formula for survival.

Some of the studies identified and included in this systematic review have more complex interventions e.g. implementation of a hospital RRT using simulation as a tool for

implementation. SBT seems to be a factor for the potential effect of rapid response system (RRS). Jung et al. (98) describes how SBT training is used in the implementation of a RRS and measures patient outcome after implementation. The authors discuss that earlier reviews have found conflicting results and K4 impact after introduction of RRS and RRT in hospitals. Sandroni et al. (113) found that different implementation strategies, educational intervention methods and training modalities might be part of the explanation why some studies find changes at K4 level, and some do not. The effect of simulation as a training modality for RRT was demonstrated by Knight et al.(99), where the probability of surviving an in-hospital cardiac arrest increased from 40,3% to 60,9% after introduction of an in situ simulation program for the resuscitation team. Theilen et al 2013 (19) found similar K4 level changes in their study, where they evaluated the K4 level effects of repeated simulation-based team training on a regular basis. The authors describe how the weekly in situ simulation trainings were designed, organized and conducted. The study demonstrated significant reduction in overall hospital mortality and highlighted the importance of regular and repeated training for team members as one of the likely reasons for their findings of correlation between in situ simulation training and K4 level changes. Theilen et al. also proved both the short term and longitudinal effects of weekly in situ simulation training (19, 88)

As presented in table 2, three studies did not find any changes at K4 level (102, 104, 105). In all three studies the intervention comprised only a one-day SBT but no repeated training. As Ersdal et al. (86) could show, a single-day intervention might have an impact on K2 and K3 level but not on K4 level. The results revealed that 18 (55%) of the studies included repeated training as part of the intervention design. All of these studies found K4 level changes on reduced mortality rates and number of adverse events.

Two of the included studies investigated the longitudinal effect of a single simulation intervention without repetition training (87, 96). These studies found that benefits from these one-day trainings declined after three months. Dillon et al. (96) found a drop in the number of patients surviving to discharge after one year with no follow up training of the hospital RRT. The study did measure decline in benefits/skills after three months with no follow up after the simulation intervention. Van de Ven et al. (87) also concluded that the beneficial effects from a one-day SBT declined after three months and recommended that repeated training should be done every three months. The combined results from these studies might implicate a possible relation between repeated training and K4 level changes. Ersdal et al. (86) and Mduma et al. (100) found that frequent recurrent training sessions (low dose, high frequency training) after

a simulation based bolus training are the key to change clinical behavior and improve patient outcome.

The relationship between repeated training, the impact on K4 level changes, indicate the importance of including repeated SBT as part of the simulation design and how this might be a key component in producing K4 level effect.

Even though no studies reported a reduced number of medical errors, there were four studies measuring protocol compliance, skill performance and number of recommended interventions performed before and after simulation training (22, 90, 106, 109). Others reported on improvement of human factor skills (found to be causing many medical errors e.g. lack of/unclear communication, loss of awareness and bad teamwork (15)). The improvements could lead to a reduced number of errors, but no direct correlations were presented in the studies included for review.

5.2.4 Quality assessment

The novel quality assessment tool described in 2.4 was used rating quality of all included studies. Studies were evaluated in adherence to seven quality indicators each indicator was weighted equally, counted as 1 of 7 possible, resulting in a score from 0 to 7.

Of the included studies,19 (58%) were rated 4/7 or higher, but only two (11%) reported in all required areas, achieving 7/7 points. In 14 (42%) the quality in describing their intervention was rated low. However, no relation was found between the general quality rating and whether a K4 level effect was observed or not as presented in table 2. The results presented in table 3 showed that seven (33%) of the studies published after 2016 had a high degree of guidelines compliance. This might indicate that that the chosen quality indicators in reporting did not reflect the quality for simulation training interventions. It could also mean that the quality of the studies adhered to the quality indicators, but did not report them. However, it highlights the necessity to implement and adhere to reporting standards in order to compare and learn from published research.

Two sub-analysis of the quality data were performed. The first investigated the possible relations between repeated training and K4 level changes. All studies with repeated training as part of the SBT design found positive K4 level results as reported and discussed in 5.2.3.

The second sub-analysis looked at the compliance to the reporting guidelines from Cheng et al. (30) for studies published in 2016 or later. Only seven of 16 studies had a rating of 4/7 or higher. This could mean that the reporting guidelines are not known to researchers. They do not seem to be widely implemented.

Only four of 33 included studies described simulation design and scenario in their article or attachments (24, 89, 103, 105). All of them achieved a high rating of quality of 4/7 or above. As there currently is no consensus or criteria defining medical simulation, insights to the methods and design become even more important to be able to understand and evaluate intervention and outcome.

5.2.5 Practical implications

The results retrieved from this systematic literature review might have some practical implications. The possibility to reduce mortality and adverse events could be used to favour SBT as a training modality compared to other forms of training. Training requires time, effort and budgets, and most organizations would like to know if all those resources are well spent. The indications of reduced mortality rates and reduced number of adverse events might increase the will to invest in SBT from a managerial point of view.

The development and use of a quality assessment tool for reporting of SBT training in this review, might be further improved and developed if applied by other researchers. This could provide new and better understanding on how different SBT designs correlates with results. This tool may also increase the comparability of SBTs.

The analysis of adherence to the guidelines showed incoherent results. This indicates needs for further implementation of tools to increase the adherence to the reporting guidelines by Cheng et al. (30). Those findings indicate a gap that is important to report to the simulation community. If SBT researchers become aware of the gap and the potential value of closing it, this might motivate for better adherence to the reporting guidelines by Cheng et al. (30)

This review found that repeated SBT led to changes in K4 level and makes the simulation community more aware of the importance of repeated training as an integrated part of the SBT design. Repeated training may be therefore be a key component in changing professional and clinical behaviour (41, 100) and for sustained effects (87, 88) of the SBT.

As discussed in the research context, the study design by the Mduma et al. were found to be an effective model demonstrating that K3 level. Observations of clinical behaviour transferred into K4 level changes. The study also demonstrates how the *The formula for survival* (70) applies to SBT related research.

The fact that no studies reported medical errors indicates a need for further research in this area. According to the WHO concerns about 4 out of 10 patients may be injured by primary or outpatient care in low- and mid-income countries. There is a need for efficient SBT interventions to reduce the risk in these healthcare systems, and there is also a need for more research investigating the educational efficiency of these SBTs, how study design correlates to K3 level changes and whether these K3 level transfers to improved survival rates.

5.3 Limitations

There are several limitations to this study.

It is uncertain whether all relevant studies have been identified and included in this review. It was a challenge to design literature search strategies and ensure that they will identify all relevant studies. Results from searches depend on how the search strategy works in different databases. Different search strategies were designed for the literature searches conducted for this review, but the number of studies identified through other sources might indicate that the strategies used did not identify all relevant studies.

Studies could also be missed because of the exclusion of non-English language. Inclusion of non-English studies could add to the evidence base and the consequences could be missing out on relevant studies that might have implications for the results and conclusions for this study.

Study design and measure of effectiveness in the included studies were heterogeneous, limiting the possibilities of more advanced analyses like meta-analyses of the results or identifying relevant factors for positive results on the K4 level besides repeated training.

The decision not to use a recognized risk of bias assessment tool could affect the results if relevant bias was not identified. We chose not to use GRADE or other tools for the assessment of bias in this review, as discussed in chapter 2.2. The included studies were already peer-reviewed and had been assessed for risk of bias.

A novel tool for quality assessment was developed for this review which could be a source for errors. A non-validated evaluation tool might introduce a risk of unprecise or missed results. The lack of correlation between quality in reporting and results on K4 level could be a result of this, but could also be the result of lack of adherence to the reporting guidelines from Cheng et al. (30).

The criteria chosen for K4 level outcome (mortality, adverse events and medical errors) might be unprecise. Despite a theoretical close relationship between the criteria described in chapter 2.3, the lack of results for changes in medical errors might indicate that the criteria could have been better or more precisely defined.

5.4 Conclusions

The results of this systematic literature review of 33 included studies indicate that simulation based training might lead to reduced mortality rates and reduced adverse events. No studies found negative effects from simulation. None of the included studies measured or reported numbers of medical errors even though there were reports of positive results that indirectly could reduce the number of medical errors.

Repeated training seemed to increase the possibility for reaching K4 level improvements. The repetition interval needed to prevent decline of skills has been found to be three months, but only in two studies. More evidence is needed to make general recommendations. There were also results implicating that there might be a relationship between repeated training and sustained effects. This study was not able to identify other design characteristics that seemed to affect K4 level changes.

Further research is needed investigating and reporting at K4 levels to be able to conclude on the best ways to perform SBT in order to maximize clinical benefits. To enable future reviews, it should be emphasized that documentation adherence to reporting guidelines as for example by Cheng et al. (30) will be of great value enabling more thorough analyzes on how intervention quality and design affects K4 level changes.

6.0 References

1. Donaldson MS, Corrigan JM, Kohn LT. To err is human: building a safer health system: National Academies Press; 2000.
2. Naik VN, Brien SE. Review article: simulation: a means to address and improve patient safety. Can J Anaesth. 2013;60(2):192-200.
3. World Health Organization. Patient safety: making health care safer. World Health Organization; 2017.
4. World Health Organization. 10 facts on patient safety 2019 [Available from: https://www.who.int/features/factfiles/patient_safety/en/.
5. Lopreiato J. Downing D, Gammon W, Lioce L, Sittner B, Slot V, Spain AE (Associate Eds.), and the Terminology & Concepts Working Group. Healthcare Simulation Dictionary, 2016. 2017.
6. Gaba DM. The future vision of simulation in health care. BMJ quality & safety. 2004;13(suppl 1):i2-i10.
7. Kirkpatrick DL, Kirkpatrick JD. Evaluating training programs . San Francisco: Barrett. Koehler Publishing, Inc; 1998.
8. Vincent C, Dr. Patient Safety, 2nd ed. 2 ed: BMJ Books; 2010 July 2010.
9. Flanagan B, Nestel D, Joseph M. Making patient safety the focus: crisis resource management in the undergraduate curriculum. Medical education. 2004;38(1):56-66.
10. World Health Organization. Patient safety curriculum guide: Multi-professional edition. 2011.
11. Helse og omsorgsdepartementet. Nasjonal helse- og sykehusplan 2020–2023. In: omsorgsdepartementet H-o, editor. 2019.
12. Torgeirsen K. Selected abstracts from the 25th Annual Meeting of the Society in Europe for Simulation Applied to Medicine. Advances in Simulation. 2019;4(2):13.
13. Graff I, Ghamari S, Schacher S, Glien P, Fimmers R, Baehner T, et al. Improvement of polytrauma management-quality inspection of a newly introduced course concept. J Eval Clin Pract. 2017;23(6):1381-6.
14. Walker D, Cohen S, Fritz J, Olvera-García M, Zelek S, Fahey J, et al. Impact Evaluation of PRONTO Mexico: a Simulation-Based Program in Obstetric and Neonatal Emergencies and Team Training. Simulation in healthcare [Internet]. 2016; 11(1):[1-9 pp.]. Available from: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-01260770/full.
15. Neily J, Mills PD, Young-Xu Y, Carney BT, West P, Berger DH, et al. Association between implementation of a medical team training program and surgical mortality. Jama. 2010;304(15):1693-700.
16. Andreatta P, Saxton E, Thompson M, Annich G. Simulation-based mock codes significantly correlate with improved pediatric patient cardiopulmonary arrest survival rates. Pediatric Critical Care Medicine. 2011;12(1):33-8.
17. Bellomo R, Goldsmith D, Uchino S, Buckmaster J, Hart G, Opdam H, et al. Prospective controlled trial of effect of medical emergency team on postoperative morbidity and mortality rates. Critical care medicine. 2004;32(4):916-21.
18. Phipps MG, Lindquist DG, McConaughey E, O'brien JA, Raker CA, Paglia MJ. Outcomes from a labor and delivery team training program with simulation component. American journal of obstetrics and gynecology. 2012;206(1):3-9.
19. Theilen U, Leonard P, Jones P, Ardill R, Weitz J, Agrawal D, et al. Regular in situ simulation training of paediatric medical emergency team improves hospital response to deteriorating patients. Resuscitation. 2013;84(2):218-22.
20. Msemo G, Massawe A, Mmbando D, Rusibamayila N, Manji K, Kidanto HL, et al. Newborn mortality and fresh stillbirth rates in Tanzania after helping babies breathe training. Pediatrics. 2013;131(2):e353-e60.

21. Riley W, Davis S, Miller K, Hansen H, Sainfort F, Sweet R. Didactic and simulation nontechnical skills team training to improve perinatal patient outcomes in a community hospital. Joint commission journal on quality and patient safety / joint commission resources. 2011;37(8):357-64.
22. Draycott TJ, Crofts JF, Ash JP, Wilson LV, Yard E, Sibanda T, et al. Improving neonatal outcome through practical shoulder dystocia training. Obstet Gynecol. 2008;112(1):14-20.
23. Lutgendorf MA, Spalding C, Drake E, Spence D, Heaton JO, Morocco KV. Multidisciplinary In Situ Simulation-Based Training as a Postpartum Hemorrhage Quality Improvement Project. Military Medicine. 2017;182(3):e1762-e6.
24. Mehta N, Boynton C, Boss L, Morris H, Tatla T. Multidisciplinary difficult airway simulation training: two year evaluation and validation of a novel training approach at a District General Hospital based in the UK. Eur Arch Otorhinolaryngol. 2013;270(1):211-7.
25. Gjeraa K, Moller TP, Ostergaard D. Efficacy of simulation-based trauma team training of non-technical skills. A systematic review. Acta Anaesthesiol Scand. 2014;58(7):775-87.
26. Merien AE, van de Ven J, Mol BW, Houterman S, Oei SG. Multidisciplinary team training in a simulation setting for acute obstetric emergencies: a systematic review. Obstet Gynecol. 2010;115(5):1021-31.
27. Boet S, Bould MD, Fung L, Qosa H, Perrier L, Tavares W, et al. Transfer of learning and patient outcome in simulated crisis resource management: a systematic review. Can J Anaesth. 2014;61(6):571-82.
28. Calvert KL, McGurgan PM, Debenham EM, Gratwick FJ, Maouris P. Emergency obstetric simulation training: how do we know where we are going, if we don't know where we have been? Australian & New Zealand Journal of Obstetrics & Gynaecology. 2013;53(6):509-16.
29. Doumouras AG, Keshet I, Nathens AB, Ahmed N, Hicks CM. A crisis of faith? A review of simulation in teaching team-based, crisis management skills to surgical trainees. J Surg Educ. 2012;69(3):274-81.
30. Cheng A, Kessler D, Mackinnon R, Chang TP, Nadkarni VM, Hunt EA, et al. Reporting guidelines for health care simulation research: extensions to the CONSORT and STROBE statements. Advances in Simulation. 2016;1(1):25.
31. Versantvoort JMD, Kleinhout MY, Ockhuijsen HDL, Bloemenkamp K, De Vries WB, Hoogen A. Helping Babies Breathe and its effects on intrapartum-related stillbirths and neonatal mortality in low-resource settings: A systematic review. Archives of Disease in Childhood. 2019.
32. Goldshtein D, Krensky C, Doshi S, Perelman VS. In situ simulation and its effects on patient outcomes: a systematic review. BMJ Simulation and Technology Enhanced Learning. 2019:bmjstel-2018-000387.
33. Lewis KA, Ricks TN, Rowin A, Ndlovu C, Goldstein L, McElvogue C. Does Simulation Training for Acute Care Nurses Improve Patient Safety Outcomes: A Systematic Review to Inform Evidence-Based Practice. Worldviews on Evidence-Based Nursing. 2019;16(5):389-96.
34. Yucel C, Hawley G, Terzioglu F, Bogossian F. The Effectiveness of Simulation-Based Team Training in Obstetrics Emergencies for Improving Technical Skills: A Systematic Review. Simulation in Healthcare. 2020;15(2):98-105.
35. Flin R, O'Connor P. Safety at the sharp end: a guide to non-technical skills: CRC Press; 2017.
36. Lee A, Bishop G, Hillman K, Daffurn K. The medical emergency team. Anaesthesia and intensive care. 1995;23(2):183-6.
37. Barrows HS, Abrahamson S. The programmed patient: a technique for appraising student performance in clinical neurology. Academic Medicine. 1964;39(8):802-5.
38. Rosen KR. The history of medical simulation. Journal of critical care. 2008;23(2):157-66.
39. Meakim C, Boese T, Decker S, Franklin AE, Gloe D, Lioce L, et al. Standards of best practice: Simulation standard I: Terminology. Clinical Simulation in Nursing. 2013;9(6):S3-S11.
40. Simulation INAfC, Learning. Standards of best practice: Simulation. Clinical Simulation in Nursing. 2011;7(4 Supplement):S1-S20.
41. Murthy NN, Challagalla GN, Vincent LH, Shervani TA. The impact of simulation training on call center agent performance: A field-based investigation. Management Science. 2008;54(2):384-99.

42. Abulebda K, Auerbach M, Limaiem F. Debriefing Techniques Utilized in Medical Simulation. StatPearls [Internet]: StatPearls Publishing; 2019.
43. Abelsson A. Learning through simulation. Disaster and Emergency Medicine Journal. 2017;2(3):125-8.
44. Kolb AY, Kolb DA. Experiential learning theory: A dynamic, holistic approach to management learning, education and development. The SAGE handbook of management learning, education and development. 2009:42-68.
45. Kolb DA. Experiential learning: Experience as the source of learning and development: FT press; 2014.
46. Bloom BS. Taxonomy of educational objectives: The classification of educational goals. Cognitive domain. 1956.
47. Sautter M, Eikeland H. The Circle of Learning in emergency medicine and healthcare education. Laerdal Medical AS; 2008.
48. Dieckmann P, Ringsted C. Pedagogy in simulation-based training in healthcare. Simulation in clinical education Oxford: John Wiley & Sons, Ltd. 2013:43-58.
49. Siemieniuk R, Guyatt G. What is GRADE? BMJ Best Practice: BMJ Best Practice; [cited 2020 09.09]. Available from: https://bestpractice.bmj.com/info/toolkit/learn-ebm/what-is-grade/.
50. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. Journal of clinical epidemiology. 2011;64(4):401-6.
51. Dictionarycom. Retrieved 01.05.2020 from http://dictionary.reference.com/browse/biasn.d.
52. Reason J. Human error: Cambridge university press; 1990.
53. Raemer D, Hannenberg A, Mullen A. Simulation safety first: an imperative. BioMed Central; 2018.
54. Bigham BL, Maher J, Morrison LJ. Patient safety in emergency medical services: advancing and aligning the culture of patient safety in EMS: Canadian Patient Safety Institute; 2010.
55. Rasmussen J. Human errors. A taxonomy for describing human malfunction in industrial installations. Journal of occupational accidents. 1982;4(2-4):311-33.
56. Hollnagel E, Braithwaite J. Resilient health care: CRC Press; 2019.
57. Hudson P. Applying the lessons of high risk industries to health care. BMJ quality & safety. 2003;12(suppl 1):i7-i12.
58. Hollnagel E, Wears RL, Braithwaite J. From Safety-I to Safety-II: a white paper. The resilient health care net: published simultaneously by the University of Southern Denmark, University of Florida, USA, and Macquarie University, Australia. 2015.
59. Bandura A, Walters RH. Social learning theory: Prentice-hall Englewood Cliffs, NJ; 1977.
60. McLeod S. Bandura-social learning theory. Retrieved from. 2011.
61. Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. Psychological review. 1993;100(3):363.
62. Sweller J. Cognitive load during problem solving: Effects on learning. Cognitive science. 1988;12(2):257-85.
63. Dekker S. The field guide to understanding'human error': CRC press; 2017.
64. Salas E, Cooke NJ, Rosen MA. On teams, teamwork, and team performance: Discoveries and developments. Human factors. 2008;50(3):540-7.
65. Baker DP, Gustafson S, Beaubien J, Salas E, Barach P. Medical teamwork and patient safety: the evidence-based relation. AHRQ publication. 2005;5(53):1-64.
66. Katzenbach JR, Smith DK. The wisdom of teams: Creating the high-performance organization. Mckinsey & Company. Inc, New York, NY. 1993.
67. Hackman JR. Why teams don't work. Leader to Leader. 1998;1998(7):24-31.
68. Katzell RA. Can we evaluate training? A summary of a one day conference for training managers. A publication of the Industrial Management Institute, University of Wisconsin. 1952(April 1952).
69. Kirkpatrick DL. How to start an objective evaluation of your training program. Journal of the American Society of Training Directors. 1956;10:18-22.

70. Søreide E, Morrison L, Hillman K, Monsieurs K, Sunde K, Zideman D, et al. The formula for survival in resuscitation. Resuscitation. 2013;84(11):1487-93.
71. Thomassen M. Vitenskap, kunnskap og praksis: innføring i vitenskapsfilosofi for helse-og sosialfag: Gyldendal akademisk; 2006.
72. Sackett DL, Rosenberg WM, Gray JM, Haynes RB, Richardson WS. Evidence based medicine: What it is and what it isn't-It's about integrating individual clinical expertise and the best external evidence. BRITISH MED JOURNAL PUBL GROUP BRITISH MED ASSOC HOUSE, TAVISTOCK SQUARE …; 1996.
73. Alper BS, Haynes RB. EBHC pyramid 5.0 for accessing preappraised evidence and guidance. Evidence Based Medicine. 2016;21(4):123-5.
74. Dixon-Woods M. Systematic reviews and qualitative methods. Qualitative research: theory, method and practice 3rd edn London: Sage. 2010:331-46.
75. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of internal medicine. 2009;151(4):264-9.
76. Fleming PS, Koletsi D, Pandis N. Blinded by PRISMA: are systematic reviewers focusing on PRISMA and ignoring other guidelines? PLoS One. 2014;9(5):e96407.
77. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Jama. 2000;283(15):2008-12.
78. Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z, et al. External validation of a measurement tool to assess systematic reviews (AMSTAR). PloS one. 2007;2(12):e1350.
79. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC medical research methodology. 2003;3(1):25.
80. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. Bmj. 2015;349:g7647.
81. Methley AM, Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. BMC health services research. 2014;14(1):579.
82. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Systematic reviews. 2016;5(1):210.
83. Olofsson H, Brolund A, Hellberg C, Silverstein R, Stenström K, Österberg M, et al. Can abstract screening workload be reduced using text mining? User experiences of the tool Rayyan. Research synthesis methods. 2017;8(3):275-80.
84. de Melo BCP, Van der Vleuten CP, Muijtjens AM, Rodrigues Falbo A, Katz L, Van Merriënboer JJ. Effects of an in situ instructional design based postpartum hemorrhage simulation training on patient outcomes: an uncontrolled before-and-after study. The Journal of Maternal-Fetal & Neonatal Medicine. 2019:1-8.
85. Van de Ven J, Van Baaren G, Fransen A, van Runnard Heimel PJ, Mol B, Oei S. Cost-effectiveness of simulation-based team training in obstetric emergencies (TOSTI study). European Journal of Obstetrics & Gynecology and Reproductive Biology. 2017;216:130-7.
86. Ersdal H, Vossius C, Bayo E, Mduma E, Perlman J, Lippert A, et al. A one-day "Helping Babies Breathe" course improves simulated performance but not clinical management of neonates. Resuscitation. 2013;84(10):1422-7.
87. van de Ven J, Fransen AF, Schuit E, van Runnard Heimel PJ, Mol BW, Oei SG. Does the effect of one-day simulation team training in obstetric emergencies decline within one year? A post-hoc analysis of a multicentre cluster randomised controlled trial. European Journal of Obstetrics & Gynecology and Reproductive Biology. 2017;216:79-84.
88. Theilen U, Fraser L, Jones P, Leonard P, Simpson D. Regular in-situ simulation training of paediatric Medical Emergency Team leads to sustained improvements in hospital response to deteriorating patients, improved outcomes in intensive care and financial savings. Resuscitation. 2017;115:61-7.

89. Ajmi SC, Advani R, Fjetland L, Kurz KD, Lindner T, Qvindesland SA, et al. Reducing door-to-needle times in stroke thrombolysis to 13 min through protocol revision and simulation training: a quality improvement project in a Norwegian stroke centre. BMJ quality & safety. 2019;28(11):939-48.
90. Arabi AM, Ibrahim SA, Manar A-R, Abdalla MS, Ahmed SE, Dempsey EP, et al. Perinatal outcomes following Helping Babies Breathe training and regular peer–peer skills practice among village midwives in Sudan. Archives of disease in childhood. 2018;103(1):24-7.
91. Bellad RM, Bang A, Carlo WA, McClure EM, Meleth S, Goco N, et al. A pre-post study of a multi-country scale up of resuscitation training of facility birth attendants: does Helping Babies Breathe training save lives? BMC pregnancy and childbirth. 2016;16(1):222.
92. Braddock 3rd CH, Szaflarski N, Forsey L, Abel L, Hernandez-Boussard T, Morton J, et al. The TRANSFORM Patient Safety Project: a microsystem approach to improving outcomes on inpatient units. JGIM: Journal of General Internal Medicine. 2015;30(4):425-33.
93. Carlo WA, Goudar SS, Jehan I, Chomba E, Tshefu A, Garces A, et al. Newborn-care training and perinatal mortality in developing countries. New England Journal of Medicine. 2010;362(7):614-23.
94. Chang OH, Levy B, Lytle H, Pope R, Phiri H, Gellhaus T, et al. Implementation of the Alliance for Innovation on Maternal Health Program to Reduce Maternal Mortality in Malawi. Obstet Gynecol. 2019;133(3):507-14.
95. Christensen B, Mora A, Benjamin B, Blough B, Duewall J, Columbus C. Baylor University Medical Center, Dallas, TXResident Training in Code Blue Execution in a Simulation Lab Improves Immediate Post-Code Survival. Ochsner Journal. 2016;16(Spec AIAMC Iss):15-6.
96. Dillon P, Moriarty H, Lipschik G. Using simulation with interprofessional team training to improve RRT/code performance. Journal of Interprofessional Education & Practice. 2018;11:67-72.
97. Goudar SS, Somannavar MS, Clark R, Lockyer JM, Revankar AP, Fidler HM, et al. Stillbirth and newborn mortality in India after helping babies breathe training. Pediatrics. 2013;131(2):e344-e52.
98. Jung B, Daurat A, De Jong A, Chanques G, Mahul M, Monnin M, et al. Rapid response team and hospital mortality in hospitalized patients. Intensive Care Medicine. 2016;42(4):494-504.
99. Knight LJ, Gabhart JM, Earnest KS, Leong KM, Anglemyer A, Franzon D. Improving code team performance and survival outcomes: implementation of pediatric resuscitation team training. Critical care medicine. 2014;42(2):243-51.
100. Mduma E, Ersdal H, Svensen E, Kidanto H, Auestad B, Perlman J. Frequent brief on-site simulation training and reduction in 24-h neonatal mortality—An educational intervention study. Resuscitation. 2015;93:1-7.
101. Sodhi K, Singla MK, Shrivastava A. Institutional resuscitation protocols: do they affect cardiopulmonary resuscitation outcomes? A 6-year study in a single tertiary-care centre. Journal of anesthesia. 2015;29(1):87-95.
102. Copson S, Calvert K, Raman P, Nathan E, Epee M. The effect of a multidisciplinary obstetric emergency team training program, the In Time course, on diagnosis to delivery interval following umbilical cord prolapse - A retrospective cohort study. Australian & New Zealand Journal of Obstetrics & Gynaecology. 2017;57(3):327-33.
103. Fransen A, van dVJ, Schuit E, van TA, Mol B, Oei S. Simulation-based team training for multi-professional obstetric care teams to improve patient outcome: a multicentre, cluster randomised controlled trial. Bjog [Internet]. 2017; 124(4):[641-50 pp.]. Available from: https://www.cochranelibrary.com/central/doi/10.1002/central/CN-01337677/full.
104. Fuhrmann L, Perner A, Klausen TW, Ostergaard D, Lippert A. The effect of multi-professional education on the recognition and outcome of patients at risk on general wards. Resuscitation. 2009;80(12):1357-60.
105. Wehbe-Janek H, Pliego J, Sheather S, Villamaria F. System-Based Interprofessional Simulation-Based Training Program Increases Awareness and Use of Rapid Response Teams. Jt Comm J Qual Patient Saf. 2013;39(8):279-87.

106. Inglis SR, Feier N, Chetiyaar JB, Naylor MH, Sumersille M, Cervellione KL, et al. Effects of shoulder dystocia training on the incidence of brachial plexus injury. American journal of obstetrics and gynecology. 2011;204(4):322. e1-. e6.
107. Nelissen E, Ersdal H, Mduma E, Evjen-Olsen B, Twisk J, Broerse J, et al. Clinical performance and patient outcome after simulation-based training in prevention and management of postpartum haemorrhage: an educational intervention study in a low-resource setting. BMC pregnancy and childbirth. 2017;17(1):301.
108. Shoushtarian M, Barnett M, McMahon F, Ferris J. Impact of introducing practical obstetric multi-professional training (PROMPT) into maternity units in victoria, Australia. BJOG: An International Journal of Obstetrics & Gynaecology. 2014;121(13):1710-8.
109. Siassakos D, Hasafa Z, Sibanda T, Fox R, Donald F, Winter C, et al. Retrospective cohort study of diagnosis–delivery interval with umbilical cord prolapse: the effect of team training. BJOG: An International Journal of Obstetrics & Gynaecology. 2009;116(8):1089-96.
110. Wagner B, Meirowitz N, Shah J, Nanda D, Reggio L, Cohen P, et al. Comprehensive perinatal safety initiative to reduce adverse obstetric events. Journal for Healthcare Quality. 2012;34(1):6-15.
111. Gaba D. Improving anesthesiologists' performance by simulating reality. 1992.
112. Vandenbroucke JP. Why do the results of randomised and observational studies differ? : British Medical Journal Publishing Group; 2011.
113. Sandroni C, Caricato A. Are rapid response systems effective in reducing hospital mortality? Critical care medicine. 2013;41(2):679-80.

7.0 Appendix

7.1 Data extraction table

Study (author/ year)	Study design		Setting	Population	Intervention	Outcome Measure?	Results K4 level
1.Ajmi et al. 2019	Before/after intervention	Prospective cohort study	Hospital, ER/ Neurology (Single center) Acute stroke patients.	Multidisciplinary: Stroke team members (n=210) Patient mimicking stroke symptoms: Patients (n= 589)	Simulation training: 26 training sessions, implementation of new protocol	Dichotomized mRS scores ('excellent', 'good' and 'worst') *, no symptoms and all-cause mortality 90 days post-stroke K3 outcome: Door to needle time reduced from 27 to 13 minutes.	Reduced 90 days mortality: 9,1 to 3,5% Patient outcome 90 days post-stroke improved. Worst outcome (deceased or bed-ridden) reduced from 12,2 to 3,5%
2.Andreatta et al. 2011	Longitudinal mixed method with control group	Prospective observational cohort study	Children's hospital (single center) Multidisciplinary	Patients n=252 Participants n=228	Repeated mock code Simulation based training program, 48 months intervention period	Survival rate = survived CPA and where discharged.	From 33% to 50% increased survival rates
3.Arabi et al. 2018	Before/after intervention	Observational prospective cohort study	Multicenter 6 centers, community midwives Obstetric/ pediatric emergency	Patients n=4390 Participants (midwives n=71)	Helping babies breathe training	24 hours survival rates: fresh stillborn (FSB) and early neonatal death (ENND) 24 hours.	Reduction from 50% death in newborns receiving mouth to mouth ventilation to 11% in newborns receiving bag-mask ventilation after HBB training.

							FSB rates decreased from 10.5 to 3.3 per 1000 births ((χ2) =8.6209, =0.003), while ENND rates decreased from 13.5 to 4.3 per 1000 live births ((χ2) =10.9369, p=0.001) pre-HBBT+RPPSP and post-HBBT+RPPSP, respectively.
4.Bellad et al. 2016	Before/after intervention	Multicenter Observational retrospective cohort study	Obstetric/ pediatric emergencies	Patients n=70,704 births for two 12-month study periods Participants =2227	Helping babies' breath training	Perinatal mortality	There were no significant differences in PMR among all registry births; however, a post-hoc analysis stratified by birthweight documented improvement in <2500 g mortality in Belgaum in both registry and in HBB trained facility births Reduction ranging

							from 46% in stillbirths to 17% in 7 days mortality. No improvement in <2500 g mortality measures was noted in Nagpur or Kenya and there was no improvement in normal birth weight survival.
5. Braddock et al. 2015	Before/after intervention	Observational prospective cohort study	Four medical and surgical inpatient units within an academic university medical center was included, with registered nurses and residents representing study participants Conducted at Stanford Hospital, a 450-bed, Level I trauma academic medical center	Patients n=13743 Participants n= 330	In situ simulation training; debriefing of medical emergencies; monthly patient safety team meetings; patient safety champion role; interdisciplinary patient safety conferences; recognition program for exemplary teamwork.	Hospital acquired sepsis, shock, unplanned higher level of care (HLOC) transfers and mortality	Rates of hospital-acquired severe sepsis/septic shock and acute respiratory failure decreased from 1.78 to 0.64 (p=0.04) and 2.44 to 0.43 (per 1,000-unit discharges) (p=0.03). Reduced incidents of acute respiratory failure 2,44 to 0,43 (per 1000-unit discharges) The mean number of days between

							cases of severe sepsis/septic shock increased from baseline to the intervention period (p=0.03). Unplanned transfers to higher level of care increased from 715 to 764 per 1,000-unit transfers (p=0.08). The weighted risk-adjusted observed-to-expected mortality ratio on all study units decreased from 0.50 to 0.40 (p<0.001).
6.Carlo et al. 2010	Cluster randomized participant groups comparison (different training modalities)	Observational prospective cohort study	Rural communities in seven sites of the Global Network for Women's and Children's Health Research in six Countries (Argentina, Democrati	Patients n= 63729 Participants: NA Birth attendants (no nurses or doctors)	A modified version of the American Academy of Pediatrics Neonatal Resuscitation Program (training resuscitation in depth).	Neonatal mortality (7 days) Stillbirths Perinatal mortality Cluster randomized training modality differences (AAPNR program or The Essential Newborn	Significant reduction in the rate of stillbirth (relative risk with training, 0.69; 95% CI, 0.54 to 0.88; P = 0.003) Otherwise no differences between the cluster

			c Republic of Congo, Guatemala, India, Pakistan, and Zambia)			Care program)	randomized groups. In subgroup analyses according to the category of birth attendant, the rate of stillbirth decreased significantly when nurses or midwives assisted the birth, but not when physicians assisted!
7.Chang et al. 2019	Before/after intervention	Observational prospective cohort study	Central hospital and a district health center in Malawi (data not collected from district center). Multidiciplinary: Full-time obstetrician–gynecologists, obstetrics and gynecology residents, medical and clinical officers, nurse midwives, nurse administrators,	Patients in total n= 2694 Participants n=128 Pre-n= 890 Education n= 759 After n= 1045	The program included classroom didactics on obstetric hemorrhage, teamwork protocols, skills laboratory activities, and simulation training	Maternal mortality from obstetric hemorrhage	In the pre- and educational period, no significant reduction of mortality was found. In the postintervention period, the rate of maternal mortality from obstetric hemorrhage decreased significantly from 1,2 to **0.2% (P=.02), which is a relative decrease of 82.1%** from the baseline preinterve

			and anesthetists who worked in labor units				ntion rate. However, confidence interval is broad for this finding, with only 8 and 2 patients in the before and after group and the numbers could be random. Data should be collected in a longer period of time or included more hospitals.
8.Christensen et al. 2016	Before/after intervention	Observational prospective cohort study	Residents at teaching institutions, in hospital cardiac arrest	Patients n=200 Participants n=21	Simulation-based code blue training program with a 3G Sim Man involving 21 internal medicine residents who were given lectures about roles/responsibilities and exposed to progressively more challenging code scenarios in which ACLS was implemented. Faculty provided feedback after each session.	Immediate post code survival and survival to discharge	Increased immediate post-code survival in the intervention cohort: 72 controls (67.3%) vs 128 intervention (71.1%) patients (*P*=0.496). This trend did not translate to increased survival to discharge: 25 controls (23.4%) vs 40 intervention (22.2%) patients (*P*=0.823)

9.Copson et al. 2017	Before/after intervention	Observational retrospective cohort study	King Edward Memorial Hospital, a tertiary referral center in Western Australia. Multidisciplinary obstetric emergency team training program	Patients n=95 Participants 100 per year (2006-2013) n= 800? (not specified, might be the same persons trained every year)		Improvement in the management of cord prolapse, in particular the diagnosis to delivery interval. We also aimed to investigate if an improvement in perinatal outcomes could be demonstrated	No significant findings for the main objective outcomes
10. Dillon et al. 2018	Before/after intervention	Observational prospective cohort study	Hospital Rapid Response Team The Corporal Michael J. Crescenz VA Medical Center (CMC VAMC)	Patients n=10354 Participants n=<200	In-situ simulation, Multidisciplinary RRT team training. The aims of the program were multifaceted: improve patient outcomes; increase healthcare team members' knowledge, competence, and confidence in emergency situations; facilitate Multidisciplinary communication and teamwork; and utilize simulation as a reality context for practice.	Patient outcome (prospective cohort) and staff outcome (mixed methods)	Significantly more patients survived to discharge after RRT training was implemented (Chi-square [13] ¼ 4.509, p < .000), but the number surviving dropped in 2012, indicating that the effect of RRT training was not maintained. Fig. 2 shows this effect.
11. Draycott et al. 2008	Before/after	Observational	Hospital, obstetric	Patients; 29025	Midwives, obstetric	Management of and	Dystocia resolution

	intervention	retrospective cohort study	staff, Southmead Hospital, Bristol	Participants; NA	staff. Emergency training, in situ The aim of this study was to compare the management of shoulder dystocia and neonatal injury associated with shoulder dystocia before and after the introduction of shoulder dystocia training for all staff in a single maternity unit.	clinical outcome/neonatal morbidity after shoulder dystocia	maneuver was used in only 49% of births complicated by shoulder dystocia, increasing significantly to 92% after training The rate of obstetric brachial plexus injury at birth was 7.0% before training and 2.3% after training. Overall injuries reduced from 9,3 – 2,3% (from 30 to 6. 24x100/30 =80%)
12. Fransen et al. 2017	Cluster randomized comparison	Multicenter Observational retrospective cohort study	Multidisciplinary Obstetric teams	Patients/births n=28657 Participants n=471	1-day simulation based training, sim center	Primary outcome was a composite outcome of obstetric complications during the first year postintervention, including low Apgar score, severe postpartum hemorrhage, trauma due to shoulder dystocia,	The composite outcome of obstetric complications did not differ between study groups [odds ratio (OR) 1.0, 95% confidence interval (CI) 0.80–1.3]. Team training reduced

						eclampsia and hypoxic-ischemic encephalopathy. Maternal and perinatal mortality was also registered.	trauma due to shoulder dystocia (OR 0.50, 95%CI 0.25–0.99) and increased invasive treatment for severe Postpartum hemorrhage (OR 2.2, 95% CI 1.2–3.9) compared with no intervention. Other outcomes did not differ between study groups
13. Fuhrman et al. 2009	Before/after intervention	Observational prospective cohort study	Multidisciplinary full-scale simulation-based education of staff on the mortality and staff awareness of patients at risk on general wards	Patients/emergencies n= 1573 Participants n= 220	1-day simulation based course.	The primary outcome measure was the rate of nursing staff awareness of patients at risk in the evening during the pre- and post-intervention periods and the secondary outcome measures were 30- and 180-day mortality and length of hospital stay for pasients at risk.	No significant differences be between intervention and control group.

14. Goudar et al. 2013	Before/after intervention	Observational prospective cohort study	Birth attendants from rural primary health centers, district and urban hospitals in southern India	Patients/births n= 9598 Participants n= 599	Helping baby's breath training Single day course + refresher training	Primary outcomes; stillbirth and neonatal mortality rate	Stillbirth declined significantly from 3.0-2,3% fresh stillbirth decreased significantly from 1,7% - 0,9% No significant increase found in neonatal mortality rates indicating that resuscitated infants survived the neonatal period.
15. Inglis et al. 2011	Before/after intervention	Retrospective observational cohort study	Multidisciplinary staff training of labor and delivery staff at urban hospital in New York, US	Patients/births n= 18677 Participants n=NA	Introduction of a simple protocol by single simulation based training intervention	Incidence of obstetric brachial plexus injury (OBPI)	Overall incidence of OBPI decreased significantly from 0,40 to 0,14 after training. OBPI after shoulder dystocia dropped significantly from 30% to 10,67%
16. Jung et al. 2016	Before/after intervention	Retrospective observational cohort study	Multidisciplinary rapid response team members Data comparison of data from 4 hospitals in urban areas in	Patients n= 161071 Participants n=NA	Introduction of rapid response teams implemented with simulation based training. Training intervention not clearly described, but was	Unexpected mortality rate (per 1000 discharges) Secondary outcomes: overall mortality, cardiac arrest rate per 1000 discharges occurring outside the	Unexpected mortality significantly decreased from 21,9 to 17,4 per 1000 discharges. Estimated to represent 1,5 saved life per week.

			Montpellier in France. Intervention at one hospital.		conducted during a 6 months period. Training intervention included	ICU and surgical wards, do not resuscitate patient deaths and length of hospital stay	Overall mortality significantly decreased from 39,9 to 34,6 per 1000 discharges.
17. Knight et al. 2014	Before/after intervention	Prospective observational cohort study	Multidisciplinary rapid response team members	Patients n=170 In situ simulations n=16 Participants not counted.	To determine whether a Composite Resuscitation Team Training program is associated with improved post-CPA survival to discharge	Improved post in-hospital cardiac arrest survival to discharge.	Post cardiac arrest survival to discharge increased from 40,3% to 60,9%
18. Mduma et al. 2015	Before/after education intervention study	Prospective observational cohort study	Multidisciplinary midwives, birth attendants	Patients n=9708 Participants n=NA	1-day HBB course followed up by a monthly 40 minutes training. Daily 3 minutes training sessions whenever time permitted (on duty).	Reduction of 24-h neonatal mortality	Reduction of 24-h neonatal mortality reduced by 40%
19. Msemo et al. 2013	Before/after education intervention	Prospective observational cohort study	Multidisciplinary Midwives, birth attendants	Patients n= 86624 Participants n=NA	1-day HBB course, refresher training and mandatory training before every shift.	Reduced 24h neonatal mortality and fresh stillbirths	Reduction of 24-h neonatal mortality by 47% and reduction of fresh stillbirth by 24%
20. Mehta et al. 2013	Before/during the intervention	Retrospective observational cohort study	Multidisciplinary	Patients n=NA Participants n=78	1-day airway management course	Reduced mortality inpatient airway fatalities	Reduced Trust mortality (announced to be re-audited due to short

							follow up period)
21. Neily et al. 2010	Before/after intervention	Retrospective observational cohort study	Multidisciplinary surgical teams, multicenter study (74 hospitals) with control group (34 facilities)	Patients n= NA Procedures n=182409 Participants n=NA	1-day simulation training integrated in a training program.	Mortality rate (30 days survival after operation) 1 year after training intervention	Observed mortality reduction 18% vs. 7% in control group. Propensity matched (risk adjusted) annual mortality reduction 50% vs. control group.
22. Nelissen et al. 2017	Before/after intervention	Prospective observational cohort study	Multidisciplinary (including ambulance drives)	Patients n= 9446 Participants n=43	½ day HMS BAB training program	Incidence of post-partum hemorrhage	38% reduction in incidence of post-partum hemorrhage
23. Phipps et al. 2012	Before/after intervention	Prospective observational cohort study	Multidisciplinary obstetric teams	Patients n= NA Participants n=186	4h simulation training as an integrated part of a 1-day (4+4h) MedTeams curriculum.	Adverse Outcome Index	Reduction of AOI score (0,052 – 0,043) statistically significant
24. Riley et al. 2011	Before/after intervention	Prospective observational cohort study	Multidisciplinary obstetric teams Multicenter (1 hospitals) with control group (2 hospital) for the simulation intervention	Patients n=NA Participants n=136 Full sim intervention participants n=36	In situ training	Perinatal morbidity	37% reduction in perinatal morbidity
25. Shoushtarian et al. 2014	Before/after intervention	Retrospective observational cohort study	Multidisciplinary obstetric teams	Patients n= 43408 Participants n=8 maternity units	PROMPT (practical obstetric multi-professional training)	Apgar score (1 and 5 minutes), cord lactate, blood loss and length of baby's	Significant improvements in cord lactate.

						stay in hospital	Length of baby' stay in hospital was reduced significantly during, but not post-training Changes in Apgar scores and number of cases with high blood loss were not significant.
26. Siassakos et al. 2009	Before/after intervention	Single center Retrospective observational cohort study	Multidisciplinary obstetric teams	Patients n=62 Participants n=NA	1-day course, annual mandatory attendance	Diagnosis – delivery interval with umbilical cord prolapse, proportion of caesarean section (CS), type of anesthesia for CS, rate of low 5-minute Apgar scores, rate of admission to neonatal intensive care unit	Significant reduction of median diagnosis – delivery interval (22-14,5 min), significant increase in the proportion of CS where recommended actions had been performed (38,46 to 82,35%),
27. Sodhi et al. 2015	Before/after intervention	Single center Prospective and retrospective observational cohort study	Code blue/in hospital cardiac arrest Multidisciplinary team training	Patients n=2164 Participants n=NA 8 simulation drills and 22 mock codes	In-situ training	Immediate survival, survival to discharge ratio, day/night survival and response time	Increased survival percentage (26,7-40,8%), survival to discharge ratio increased from 23,4 to 66,6%), day/night survival improved, and response

							time improved from 4 to 1,5 min.
28. Theilen et al. 2013	Before/after training	Single center Prospective cohort study	Multidisciplinary team training, pMET teams.	Patients n=16506 Participants n= 6 participants per weekly in-situ SBT	In situ training	Recognition of deterioratin g patients,	1.Deteriorating patients were recognized more promptly (4 to 1,5h) improvement. 2.More often reviewed by consultants (45 to 76%) 3.Associated PICU admissions 4.Reduced (56-51, p=0,02) and PICU bed days (527 to 336)
29. Theilen et al. 2017	Longitudinal effect, 3 years after intervention compared to 1-year results	Single center Prospective observational cohort study	Long term impact of regular training.	Patients n=17096 Participants n= 6 participants per weekly in-situ SBT	In situ training	Same as the above	3-year reduction of 1. 0,5h 2. 81% 3. 32 4. 193
30. Van de Ven et al.(I) 2017	Longitudinal effect of training 1 year after training	Prospective observational cohort study Cluster randomized and controlled trial	Multidisciplinary obstetric teams Multicentre study	Patients n=29063 Participants n= 471 team members from 24 obstetric units (randomized)	1-day team training	Neonatal and maternal obstetric complications; Composite outcome of low Apgar score, severe postpartum hemorrhage, trauma due to shoulder dystocia,	Significant effects on trauma due to shoulder dystocia and for invasive treatment for severe postpartum hemorrhage. Effects found in

						eclampsia and hypoxic-ischemic encephalopathy	one of four quarters. Decline in effect of team training not significant (p=0,052)
31. Wagner et al. 2012	Before/after training	Prospective observational cohort study	Multidisciplinary obstetric teams	Patients n=10327 Participants n=NA	Obstetric emergency simulation program	Adverse outcome index (MAOI)	MAOI decreased significantly (60% from 2% to 0,8%). Maintained over the 2-year period. Reduced rates of return to the operating room and birth trauma.
32. Walker et al. 2016	Before/after training	Randomized controlled trial	Multidisciplinary PRONTO obstetric and neonatal emergency team training, in situ simulation, multicenter study	Patients n=51086 Participants n= 24 hospitals (12 intervention and 12 in control), 450 participants.	1-day simulation training in addition to 2 days of didactic lectures	Perinatal mortality at 12 months follow up, secondary outcome: Obstetric hemorrhage mortality, preeclampsia/eclampsia mortality and maternal complications	Incidence of hospital-based neonatal mortality was significantly lower; 40% eight months post intervention in intervention hospitals vs. control after adjustment for baseline differences, although no significant differences after 4 and 12 months. Cesarean delivery

							incidence was significantly lower; 18 to 21% in the intervention group at all time intervals
33. Wehbe et al. 2013	Before/after training	Prospective Observational Cohort trial		Patients n=79937 Participants n=359 (10 hospital units)	3 x 3-hour training sessions (1 didactic lecture and 2 simulation scenarios	Relationship of increased activation of rapid response teams and changes in mortality rates. Measuring hospital mortality rates	No significant improvement was found.

7.2 Literature searches

7.2.1 Google Scholar:

https://scholar.google.com/scholar?hl=no&as_sdt=1%2C5&as_ylo=2015&as_vis=1&q=medical+emergency+team%7CHEMS%7Crapid+response+team+simulation+training%7Cscenario+based+training+mortality%7Cmortality+rate+improved+mortality%7Creduced+mortality&btnG=

7.2.2 OVID EMBASE

Database: Embase <1974 to 2019 October 14> Search Strategy:

---

1 (emergency medical team* or medical emergency team* or HEMS or rapid response team* or patient care team* or air ambulance* or emergency helicopt* or airway response team*).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (5683)

2 (medical emergenc* or medical emergency situation* or critically sick or critically injured or trauma patient* or cardiac arrest or airway emergenc* or post-partum bleeding or stillbirth* or massive bleeding or life threatening medical situation*).mp. [mp=title, abstract,

heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (109811)

3 1 or 2 (113849)

4 ((simulation adj2 training) or patient simulation).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word] (9328)

5 simulation training/ (3884)

6 4 or 5 (9328)

7 3 and 6 (525)

8 mortality/ or mortality rate/ or (mortality or reduced mortality or improved mortality rate).mp. (1391016)

9 7 and 8 (56)

10 improved mortality rates.mp. (116)

11 mortality rate/ (45404)

12 7 and 10 (0)

****

7..2.3 OVID MEDLINE:

Search for: from 11 [(emergency medical team* or medical emergency team* or HEMS or rapid response team* or patient care
team* or air ambulance* or emergency helicopt* or airway response team* or medical emergenc* or medical emergency
situation* or critically sick or critically injured or trauma patient* or cardiac arrest or airway emergenc* or post-partum bleeding or stillbirth* or massive bleeding or life threatening medical situation*).mp. and ((((simulation
adj3 training) or (patient adj3 simulation*)).mp. or simulation training/ or scenario based training.mp.) not (computer simulation*.mp. or computer simulation/)) and (mortality/ or mortality rate/ or (mortality or ((improved or reduc*) adj3 (mortality adj3 rat*))).mp.) and (((systematic* or literature) adj3 (overview* or review* or search*)) or meta analys*).ti,ab. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol

supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]] keep 1-4

7.2.4 Epistemonikos:

(title:((title:(simulation training OR scenario based training) OR abstract:(simulation training OR scenario based training)) NOT (title:(computer simulation) OR abstract:(computer simulation*))) OR abstract:((title:(simulation training OR scenario based training) OR abstract:(simulation training OR scenario based training)) NOT (title:(computer simulation*) OR abstract:(computer simulation*))))*

7.2.5 PubMed:

Recent queries in pubmed

Search,Query,Items found,Time

"#14,""Select 7 document(s)"",7,08:36:39"

"#13,""Search (((((((emergency medical team* or medical emergency team* or HEMS or rapid response team* or patient care team* or air ambulance* or emergency helicopt* or airway response team*[tiab] or medical emergenc* or medical emergency situation* or critically sick or critically injured or trauma patient* or cardiac arrest or airway emergenc* or post-partum bleeding or stillbirth* or massive bleeding or life threatening medical situation*)))) AND ((((simulation training OR patient simulation* or scenario based training))) NOT (("""""computer simulation"""""[MeSH Terms]) AND computer simulation[Title/Abstract]))) AND (((mortality[MeSH Terms]) OR mortality rate[MeSH Terms]) OR (mortality[Title/Abstract] OR imporved mortality[Title/Abstract] OR reduced mortality[Title/Abstract])))) AND (((systematic overview*[Title/Abstract] OR systematic review*[Title/Abstract] OR systematic search*[Title/Abstract] OR literature overview*[Title/Abstract] OR literature review*[Title/Abstract] OR literature search[Title/Abstract] OR meta analys*[Title/Abstract]))"",7,08:36:10"

"#12,""Search (#11) AND (((systematic overview*[Title/Abstract] OR systematic review*[Title/Abstract] OR systematic search*[Title/Abstract] OR literature overview*[Title/Abstract] OR literature review*[Title/Abstract] OR literature search[Title/Abstract] OR meta analys*[Title/Abstract]))"",0,08:35:09"

"#11,""Select 12 document(s)"",12,08:34:53"

"#10,""Search (((((((emergency medical team* or medical emergency team* or HEMS or rapid response team* or patient care team* or air ambulance* or emergency helicopt* or airway response team*[tiab] or medical emergenc* or medical emergency situation* or critically sick or critically injured or trauma patient* or cardiac arrest or airway emergenc* or post-partum bleeding or stillbirth* or massive bleeding or life threatening medical situation*)))) AND ((((simulation training OR patient simulation* or scenario based training))) NOT ((""""""computer simulation""""""[MeSH Terms]) AND computer simulation[Title/Abstract]))) AND (((mortality[MeSH Terms]) OR mortality rate[MeSH Terms]) OR (mortality[Title/Abstract] OR imporved mortality[Title/Abstract] OR reduced mortality[Title/Abstract]))))) AND (((medical error[MeSH Terms]) OR error[MeSH Terms]) OR (error*[Title/Abstract] OR mistake*[Title/Abstract]))"",12,08:34:11"

"#9,""Search (((((emergency medical team* or medical emergency team* or HEMS or rapid response team* or patient care team* or air ambulance* or emergency helicopt* or airway response team*[tiab] or medical emergenc* or medical emergency situation* or critically sick or critically injured or trauma patient* or cardiac arrest or airway emergenc* or post-partum bleeding or stillbirth* or massive bleeding or life threatening medical situation*)))) AND ((((simulation training OR patient simulation* or scenario based training))) NOT ((""""""computer simulation""""""[MeSH Terms]) AND computer simulation[Title/Abstract]))) AND (((mortality[MeSH Terms]) OR mortality rate[MeSH Terms]) OR (mortality[Title/Abstract] OR imporved mortality[Title/Abstract] OR reduced mortality[Title/Abstract]))"",112,08:33:54"

"#8,""Search ((systematic overview*[Title/Abstract] OR systematic review*[Title/Abstract] OR systematic search*[Title/Abstract] OR literature overview*[Title/Abstract] OR literature review*[Title/Abstract] OR literature search[Title/Abstract] OR meta analys*[Title/Abstract])"",348640,08:33:14"

"#7,""Search ((medical error[MeSH Terms]) OR error[MeSH Terms]) OR (error*[Title/Abstract] OR mistake*[Title/Abstract])"",398087,08:31:35"

"#6,""Search ((mortality[MeSH Terms]) OR mortality rate[MeSH Terms]) OR (mortality[Title/Abstract] OR imporved mortality[Title/Abstract] OR reduced mortality[Title/Abstract])"",951402,08:30:30"

"#4,""Search (((simulation training OR patient simulation* or scenario based training))) NOT ((""""""computer simulation""""""[MeSH Terms]) AND computer simulation[Title/Abstract])"",23857,08:27:11"

"#3,""Search (""""""computer simulation""""""[MeSH Terms]) AND computer simulation[Title/Abstract]"",5819,08:26:55"

"#2,""Search (simulation training OR patient simulation* or scenario based training)"",24037,08:25:48"

"#1,""Search ((emergency medical team* or medical emergency team* or HEMS or rapid response team* or patient care team* or air ambulance* or emergency helicopt* or airway response team*[tiab] or medical emergenc* or medical emergency situation* or critically sick or critically injured or trauma patient* or cardiac arrest or airway emergenc* or post-partum bleeding or stillbirth* or massive bleeding or life threatening medical situation*))"",203637,08:23:32"

7.2.6 Cochrane

Search Name: Torgeirsen

Date Run: 30/10/2019 09:47:53

Comment:

ID Search Hits

#1 (emergency medical team* or medical emergency team* or HEMS or rapid response team* or patient care team* or air ambulance* or emergency helicopt* or airway response team* or medical emergenc* or medical emergency situation* or critically sick or critically injured or trauma patient* or cardiac arrest or airway emergenc* or post NEXT partum bleeding or stillbirth* or massive bleeding or life threatening medical situation*) 41092

#2 (simulation NEAR/3 training):ti,ab,kw OR (patient NEAR/3 simulation*):ti,ab,kw OR (simulation based training):ti,ab,kw NOT (computer simulation*):ti,ab,kw 2308

#3 mortality or mortality rate 90621

#4 MeSH descriptor: [Mortality] this term only 495

#5 (improved or reduc*) NEAR/3 (mortality NEAR/3 rat*) 846

#6 MeSH descriptor: [Medical Errors] explode all trees 2887

#7 (medical error* OR error* OR mistake*):ti,ab,kw 19046

#8 (((systematic* or literature) adj3 (overview* or review* or search*)) or meta analys*):ti,ab,kw 20899

#9 #1 AND #2 483

#10 #3 OR #4 OR #5 90621

#11 #9 AND #10 32

#12 #11 AND #8 3

#13 #6 OR #7 21140

#14 #11 AND #13 6

7.2.7 OVID CINHAL

NY

EBSCOhost

Thursday, October 31, 2019 7:47:00 AM

#	Query	Limiters/Expanders	Last Run Via	Results
S6	S1 AND S3	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL with Full Text	0
S5	S1 AND S2	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL with Full Text	1
S4	S1 AND S2 AND S3	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL with Full Text	0
S3	TI ((((systematic* or literature) adj3 (overview* or review* or search*)) or meta analys*)) OR AB ((((systematic* or literature) adj3 (overview* or review* or search*)) or meta analys*))	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL with Full Text	59,644
S2	(MH "Treatment Errors+") AND error* or mistake*	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL with Full Text	26,020
S1	(emergency medical team* or medical emergency team* or HEMS or rapid response team* or patient care team* or air ambulance* or emergency helicopt* or airway response team* or	Search modes - Boolean/Phrase	Interface - EBSCOhost Research Databases Search Screen - Advanced Search Database - CINAHL with Full Text	24

	medical emergenc* or medical emergency situation* or critically sick or critically injured or trauma patient* or cardiac arrest or airway emergenc* or post-partum bleeding or stillbirth* or massive bleeding or life threatening medical situation*) AND (((simulation N2 training) or (patient N2 simulation*) OR (MH simulation training+) OR (scenario based training)) NOT (computer simulation* or (MH computer simulation+))) AND ((MH mortality+ or MH mortality rate+) OR (mortality OR ((improved or reduc*) N2 (mortality N2 rat*))))