UiT The Arctic University of Norway

Faculty of Health Sciences: The Department of Community Medicine

The effectiveness of closed-loop communication training for healthcare providers on use, medical errors, and task performance: a systematic review

Maria Celine Aas

HEL-3950: Master`s Thesis in Public Health, Spring 2024

Supervisor: Professor Rigmor Berg, *Arctic University of Norway*

Table of Contents

Acknowledgements

I would like to give a special thanks to several of the employees at the Arctic University of Norway. To my supervisor, Professor Rigmor Berg, for providing me with valuable, thorough, and quick feedback, and motivating me throughout the entire process of conducting this systematic review. To Doctoral Research Fellow, Bahar Kucuk, and Associate Professor, Kristin Borch, for the many advice given during the planning phase and development of the protocol. I could not have hoped for more exceptional professional guidance.

Special mention goes to my close friends, and especially my two roommates, Ineke and Aurora, for continuously checking up on my energy levels and progress. My appreciation further extends to my colleagues and workplace, for always lending me an office space when I felt the need for social company. Gratitude must also be expressed for the time my father put into reading the review, and to the rest of my family members for the immense support received during the last six months.

Last but not least, I am endlessly grateful for the collective experience and knowledge gained from this Master`s degree in Public Health and my Bachelor`s degree in nursing. This combination has given me a comprehensive understanding of today's healthcare system, made me aware of various areas of improvement, and thereby inspired me to delve deeper into the topic of patient safety and interprofessional communication.

Abstract

Background: Patient harm from unsafe healthcare resulting in medical errors is an increasing global health challenge, ranking among the top causes of mortality on a global basis. These incidents often stem from clinical process failures, with a substantial proportion directly attributed to miscommunication and communication breakdowns. Numerous healthcare providers have admitted to having caused harm to patients due to communication shortfalls and uttered a desire to improve their communication skills. Closed-loop communication (CLC) is a long-standing skill taught in military settings to avoid miscommunication among team members, and places strong emphasis on verification to ensure that messages or orders are understood correctly. Nevertheless, CLC is still underutilized in medical practice.

Objective: The purpose of this systematic review was to investigate the effectiveness of CLC training for healthcare providers on increasing the use of CLC, reducing medical errors, and improving task performance.

Methods: Searches were performed in MEDLINE, EMBASE, CINAHL, CENTRAL, ERIC, and Google Scholar between the 10th and 14th of January 2024. To uncover further relevant studies, the reference lists of all included studies, related reviews, and relevant excluded studies, were screened. Participants eligible for inclusion were healthcare providers directly involved in patient care or assisted in providing it. CLC training in this review was understood as instructions that are designed to help individuals improve their ability to communicate effectively, with an emphasis on verification to ensure that messages or orders are understood correctly. The training had to be implemented either for the purpose of increasing the use of closed-loop communication between healthcare providers, or to achieve other desired outcomes related to patient safety through increased use of this particular communication strategy. Only randomized controlled trials were included, and risk of bias was assessed using the Cochrane (RoB 1.0) tool. The primary outcome was the frequency of closed-loop communication, and the secondary outcomes were medical errors and task performance. GRADE was used for rating the certainty of evidence. Due to considerable heterogeneity, a narrative approach was used to synthesize the results.

Results: All searches combined resulted in 1493 records, of which four studies with a total of 197 participants were included. Three studies from the U.S. and one study from Switzerland.

One study had low or unclear risk of bias for all domains, while all the remaining studies had high risk of bias for at least one domain. All studies implemented a closed-loop communication training method comparing blindfolded simulations versus non-blindfolded simulations and measured the frequency of CLC. The results from the majority of the studies suggested more frequent use of CLC in the blindfolded group compared to the non-blindfolded group. However, the certainty of evidence for the primary outcome was rated as very low. Furthermore, one of the included studies reported a non-significant difference in the frequency of CLC, thus indicating some inconsistency in the findings. Two studies found that the blindfolded simulation training significantly reduced communication errors. There was a lack of statistically significant improvements in task performance measures, apart from some indication of a positive effect on non-clinical skills among team leaders.

Conclusion: Due to the several limitations of the evidence provided, the overall findings are not convincing enough to conclude that implementing a blindfolded simulation training as closed-loop communication training for healthcare providers will guarantee increased use of CLC, a reduction in the number medical errors, improved task performance, or produce the same results as shown in this systematic review, when replicated. That being said, the results suggested a potential positive effect on communicative behavior that should not be overlooked. To best inform about the effectiveness of closed-loop communication training, more research and larger studies are needed that also ideally include measure long-term effectiveness and retention.

Abbreviations

CLC	Closed-loop communication
AHRQ	Agency for Healthcare Research and Quality
WHO	World Health Organization
PRISMA	The Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	Randomized controlled trials
RoB	Risk of Bias
SWiM	Synthesis without meta-analysis
GRADE	The Grading of Recommendations Assessment, Development and Evaluation
EM	Emergency medicine
CRM	Crisis Resource Management
GRS	Global Rating Scale
TLX	Task Load Index
CPR	Cardio-pulmonary resuscitation
N	Number
TP	Task performance
ME	Medical error
FoCLC	Frequency of closed-loop communication
SS	Simulation scenario
NI	No information
OR	Odds Ratio
CI	Confidence interval
SA	Summary assessment
OO	Objective outcomes
SE	Standard Error
SRP	Self-reported performance
IQR	Interquartile Range

List of tables

List of Figures

1 Background

1.1 Description of the problem

Patient harm from unsafe healthcare resulting in medical errors is an increasing global health challenge, ranking among the top causes of mortality on a global basis (World Health Organization [WHO], 2021). Such errors can be defined as "an act of omission or commission in planning or execution that contributes or could contribute to an unintended result" (Grober & Bohnen, 2005, p. 42). Although many medical errors do not cause serious consequences for patients, they are still found to be the third leading cause of death in the United States (Makary & Daniel, 2016). These incidents often stem from clinical process failures, with a substantial proportion directly attributed to miscommunication and communication breakdowns (WHO, 2021; Makary & Daniel, 2016). Based on this evidence, suggesting that the absence of effective communication may lead to both further mishaps during healthcare delivery and more patient harm, failing to practice communication strategies intended to protect patients can also be considered a medical error in itself.

Furthermore, numerous healthcare providers have admitted to having caused harm to patients due to communication shortfalls and uttered a desire to improve their communication skills (Zimmer et al., 2021). This underscores both the motivation to learn among healthcare providers, and the urgent need for such measures to avoid miscommunication, with the potential of averting millions of adverse events (WHO, 2021). Research suggests that interprofessional communication skills can be significantly improved through training methods such as simulations (Foronda et al., 2016). Nevertheless, the persistent lack of emphasis on patient safety during healthcare providers' education results in insufficient attention being paid to the importance of teamwork and communication in protecting patients from harm (WHO, 2021).

1.2 Description of closed-loop communication and how it might work

Closed-loop communication (CLC) is a long-standing skill taught in military settings to avoid miscommunication among team members, and places strong emphasis on verification to ensure that messages or orders are understood correctly (Burke et al., 2004). This type of communication is obtained if three key components are completed during a communication event: "(1) sender initiates message (2) receiver accepts message and provides feedback confirmation (3) sender verifies message was received" (Agency for Healthcare Research and Quality [AHRQ], 2023, p. 6). As displayed in AHRQ (2023, p. 7), a CLC event may be exemplified as follows:

Dr. Moss: *"Mary, please share the information pamphlet on cholesterol management with Mr. Garcia and arrange for him to come for a follow-up visit in a month."*

Mary: *"Confirmed. I`ll share the information pamphlet on cholesterol management and arrange a follow-up visit for Mr. Garcia in a month."*

Dr. Moss: *"Correct."*

1.3 Why it is important to do this review

Despite its absence in healthcare literature until 1995, CLC has progressively gained recognition as an invaluable tool to transfer into intricate and interdisciplinary medical settings, where the risk of miscommunication is high (Peyre, 2014; Burke et al., 2004). CLC is currently recommended as a guideline for communication in TeamSTEPPS; a well-known evidence-based framework and guide to optimize teamwork skills and performance, to ensure safety for patients (AHRQ Quality, 2023 Nevertheless, CLC is still underutilized in medical practice. Results from studies conducted by El-Shafy et al. (2018) and Härgestam et al. (2013) revealed that a mere 26.1% and 14% of all verbal orders among trauma team members, respectively, adhered to CLC principles. Another point of concern is research indicating that healthcare providers tend to consider themselves as good communicators; although they may not necessarily master CLC (Zimmer et al., 2021).

An initial search confirms that primary studies have investigated various training methods to encourage healthcare providers to use CLC more frequently, in addition to whether the utilization of this communication strategy may reduce medical errors and improve task performance. Summarizing and evaluating the existing empirical evidence on the effectiveness of CLC training may contribute towards achieving one of the strategic goals outlined in the Global Patient Safety Action Plan 2021-2030 (WHO, 2021): to deliver and design safer healthcare systems. However, such actions have not yet been carried out. Consequently, health care systems might inadvertently overlook an area in need of more indepth exploration or miss out on effective training solutions for strengthening patient safety through more effective communication.

1.4 Objective

The purpose of this systematic review was to identify, evaluate, and summarize all research evidence from studies investigating the effectiveness of CLC training for healthcare providers on increasing in the use of CLC, reducing medical errors, and improving task performance.

1.5 Research question

What is the effectiveness of closed-loop communication training for healthcare providers on use, medical errors, and task performance?

2 Methods

This systematic review was conducted in accordance with the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al., 2023). The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 checklist was also used as a reporting guideline (Page et al., 2021). Furthermore, the entire process, including the methodological parts, was undertaken exclusively by the sole author of this paper, complying with the requirements for a master`s thesis. The pre-made protocol was published in PROSPERO under the name of the author of this review, and may be found using the following registration number: CRD42024497122 (Centre for Reviews and Dissemination, n.d.)

2.1 Inclusion/Exclusion criteria

2.1.1 Population

Participants eligible for inclusion were healthcare providers directly involved in patient care or assisted in providing it. This included healthcare workers and health professionals, also including those undergoing postgraduate or specialized training. Healthcare students were also included since they commonly engage in patient care and acquire responsibility for it during their educational programs. Non-clinical personnel, including managerial, administrative and research roles were not considered. Furthermore, exclusion was applied to those involved in non-human fields of healthcare, such as veterinary and aquatic medicine. If occupational or educational roles stood unspecified, or studies included a mix of populations who did and did not meet the inclusion criteria, at least 80% of the population had to fit the inclusion criteria for studies to be included.

2.1.2 Intervention

Closed-loop communication training in this review was understood as instructions that are designed to help individuals improve their ability to communicate effectively, with an emphasis on verification to ensure that messages or orders are understood correctly. It needed to be implemented either for the purpose of increasing the use of closed-loop communication between healthcare providers, or to achieve other desired outcomes related to patient safety through increased use of this particular communication strategy. Studies were eligible for inclusion regardless of whether the training was applied under real clinical or simulated conditions. Training could be delivered virtually, in-person, on group-level, individually, or through a combination of these approaches. Courses, workshops, debriefs, simulations, educational or teaching programs, were all considered relevant training methods. Studies focusing solely on the implementation of new policies, systems, guidelines, checklists, procedures, protocols, or digital devices, without any complementary communication training, were not deemed eligible for inclusion. Beyond the above criteria, no exclusions were made based on trainer credentials, training content, nor based on the length of individual training sessions or the overall training period.

2.1.3 Comparison

Studies had to incorporate one of the following comparison groups: standard communication training, no training, another type of intervention.

2.1.4 Outcomes

The original plan was that all eligible studies had to measure at least one outcome linked to the use of closed-loop communication, and the type of outcome would not restrict inclusion. This decision was made based on the initial expectation that all such outcomes would capture the direct effect of closed-loop communication training on actual communication behaviors. Nevertheless, the broad inclusion criteria turned out to also cover outcomes that measured an indirect effect. Therefore, after careful consideration, the primary outcome was narrowed to solely include the frequency of closed-loop communication or message/order verification among healthcare providers, and could be measured using either self-reported or observed data.

Secondary outcomes of interest included observed differences or changes in all types of medical errors or task performance, the latter referring to the effectiveness with which healthcare providers execute their tasks, which conforms to the term`s definition by Broman & Motowidlo (1997). These were considered important due to their representation of a potential indirect effect of CLC training, and the ultimate value of the intervention on direct clinical patient care, in addition to the effectiveness and quality of it. Beyond the criteria above, all approaches to measuring outcomes were accepted.

2.1.5 Study designs

This systematic review included only randomized controlled trials (RCTs). Non-randomized controlled trials, controlled before-and-after studies, interrupted-time-series, and uncontrolled before-and-after studies would also have been considered if there was a lack of eligible RCTs.

2.1.6 Other criteria and reporting characteristics

Geographical location did not pose any restrictions on inclusion. Primary studies published in scientific journals between 1995 and 2024, and available in full text, were included. The broad timespan was set to cover all relevant studies published on closed-loop communication training since this communication approach became a familiar concept in health literature. Due to time constraints, unpublished studies, conference abstracts and papers, book chapters, secondary research articles, and theses were not searched for. Publications only available in languages other than English, Norwegian, Swedish, or Danish were excluded. This was because of the lack of resources for translation.

2.2 Information sources and search strategy

The decision to adopt a broader search strategy was influenced by several factors. First, defining specific search terms for the population, comparison group, and outcomes became challenging due to the inclusive eligibility criteria, especially without the risk of excluding relevant studies. Furthremore, closed-loop communication is a rather niche topic, and the number of available studies addressing this concept remains limited. Therefore, to ensure a balanced search considering both sensitivity and precision, only terms related to closed-loop communication and terms commonly used interchangeably with the concept were included, connected with Boolean “OR” and “AND” operators. Moreover, proximity operators, phrase searching, truncations, and proximity operators were used as suitable.

The literature searches were adapted to, and performed in, each of the following electronic databases: MEDLINE, EMBASE, CINAHL, CENTRAL and ERIC, between the 10th and 14th of January 2024 (inserted exactly as run in Supplementary appendix 1). As per the eligibility criteria, all searches had a date coverage from 1995 up to and including 2024. To uncover further relevant studies not identified in the chosen databases, an additional search was conducted in Google Scholar on the 10th of January 2024. Additionally, the reference lists of all included studies, related reviews, and relevant excluded studies, were screened.

2.3 Study records

2.3.1 Data management and selection process

All the identified records from database searches were stored and organized in Endnote 21 (The Endnote Team, 2013). Here, duplicates were removed by setting duplication preferences to author, year, and title, without any manual assessment. A manual check for duplicates not identified by the software was also conducted by sorting the library alphabetically first based on title, and then author. The remaining references were exported into Rayyan, where duplicates not found in Endnote were identified using the *detect duplicates* function and removed (Ouzzani et al., 2016). Furthermore, the process of screening and selecting studies for inclusion were also executed in Rayyan.

First, all titles and abstracts were reviewed in accordance with inclusion/exclusion criteria to eliminate studies that were clearly not relevant. Thereafter, for the studies not excluded thus far, full-text papers were retrieved and assessed based on the same criteria. Screening questions were prepared in advance and used to ensure consistency. The screening questions for title and abstract screening were brief and appropriate for an effective screening process, while more detailed for the full-text reading (as shown in Supplementary appendix 2). Studies excluded during full-text reading were added to a "characteristics of excluded studies" table with reasons for exclusion (see Supplementary appendix 3). The exact same full-text screening and selection process was completed for the records identified from screening reference lists.

The first step taken in scenarios where it was difficult to judge with certainty whether a study met inclusion/exclusion criteria, was to reread and reassess the study. If uncertainties remained, supervisor (Rigmor Berg) was consulted for a second opinion. The full study selection process, including the number of papers identified, screened, included, excluded, along with reasoning for the decisions made is presented in **Figure 1** using the PRISMA 2020 flow diagram (Page et al., 2021).

2.3.2 Data extraction

A standardized extraction form was designed to (1) extract data that directly addressed the research question, objective and aligned with the specific outcomes of interest, and (2) to ensure consistency in data extraction across the studies included. The following information was extracted: reference information (title, author, publication year, source), study details (country, study design), population description (sample size, occupational or educational roles), intervention description (approach, duration, simulated or real conditions), comparison description (standard training, another intervention, or no intervention), and study results (outcome data and measurement methods).

The corresponding authors of three studies included in this systematic review were contacted by email, requesting additional data needed to provide a more comprehensive analysis (Buyck et al., 2019; Lopez de Alda et al., 2021; Hughes et al., 2020). Buyck et al. (2019) replied that the requested information would be sent by the end of April at the latest. Yet, no email with data was received within the submission deadline for this master`s thesis. The remaining authors did not respond to the e-mail. A brief summary of the included studies is presented in **Table 1**, while the complete data extraction is available in the "Characteristics of included studies" table in Supplementary appendix 4.

2.4 Risk of bias assessment

The risk of bias assessment of the included RCTs was conducted using the Cochrane Risk of Bias (RoB) 1.0 tool found in the Cochrane Handbook version 5.1.0 (Higgins & Green, 2011). The updated RoB 2.0 version has been criticized for being challenging to implement even for highly experienced assessors (Minozzi et al., 2022a; Crocker et al., 2021; Sterne et al., 2019). Also, there are currently no requirements for review authors to use the revised tool, and many still adhere to RoB 1.0 instead of RoB 2.0 (Minozzi et al., 2022b). For these reasons, the original version of the tool was chosen for this systematic review. RoB 1.0 covers six sources of bias, and for each domain studies were judged to have either low risk, unclear risk, or high risk of bias.

Both risk of attrition bias and detection bias were assessed on outcome level, under the following grouping of outcomes: self-reported performance and objective outcomes. Risk of performance bias was considered to be equal for all outcomes, bearing in mind that all outcomes included in this review were behavioral-based. Therefore, performance bias was assessed using a study-level approach, similarly to selection bias and reporting bias.

In addition, a summary assessment of risk of bias across domains was conducted separately for self-reported performance and objective outcomes for each of the individual studies. Determining an overall risk of bias was mainly important for its subsequent use in data synthesis.

2.5 Data synthesis

As outlined in the protocol, the plan was to use RevMan for the purpose of conducting meta-analysis on the condition that combining statistical results from studies was considered feasible and justifiable. The combined effect estimate measure, and statistical method used to summarize and compare results, would depend on the effect measures and the type of data (e.g., dichotomous, or continuous) extracted from the included studies. A corresponding 95% confidence levels would nevertheless be provided alongside the estimate if ratio/difference were to be calculated. This quantitative synthesis would be presented in a graphical display, including a forest plot distributing the effect measures from the individual studies, a chi-squared (X^2), Tau-squared (T^2) and I-squared (I^2) statistic for measuring heterogeneity of intervention effects. The meta-analysis would be undertaken using the random-effects model and a funnel plot would be presented for assessing publication bias if the condition of at least ten studies included were met.

The above-mentioned approach would further have required that multiple studies were homogenous and consistent both in the reporting of numerical data (e.g., sample size, direction of effect, and confidence intervals), and in regard to the effect measures used. Moreover, discrepancies between the included studies in terms of population, intervention, outcome domain, setting, study design, and risk of bias should be kept to a minimum.

An alternate approach if a meta-analysis could not be carried out was to conduct a narrative synthesis, adhering to the Synthesis without meta-analysis (SWiM) reporting guidelines

(Campbell et al., 2020). This would involve grouping studies into subpanels and categorizing them based on their nature of intervention, and ordering studies according to their risk of bias assessment, unless other structuring methods seemed more suitable due to the characteristics of the included studies. As explained and justified in the results section, a narrative approach to synthesize results was ultimately used. Furthermore, after delving deeper into the studies, it turned out that a different tactic for grouping outcomes might in fact be more appropriate. Instead of the original plan, studies were grouped on the basis of how participants were organized in group settings during simulations.

The protocol also specified that, if sensible, a sensitivity analysis and subgroup analysis should be performed to address any variability and to evaluate the robustness of findings. Also, to convey findings more visually, supplementary figures should be considered (e.g., forest plot without the pooled estimate effect). Yet, due to the small number of included studies and the considerable heterogeneity between them, these steps were not acted upon.

2.6 Confidence in cumulative evidence

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) (Schünemann et al., 2013) framework was used to rate the certainty of evidence in effect estimates reported on the primary outcome, frequency of closed-loop communication, from the individual studies, grading them as either very low, low, moderate, or high certainty. All four studies started at high certainty because of their RCT study design. Therefore, the level of certainty could only be downgraded through the consideration of five domains (risk of bias, inconsistency, indirectness, imprecise results, and publication bias). If other study designs had also been included, these could potentially have been upgraded by considering further three domains: large magnitude of effect, a dose-response relationship, and all plausible confounding. Since meta-analysis was not conducted, a modified approach to apply the constructs of GRADE by Murad et al. (2017) was used.

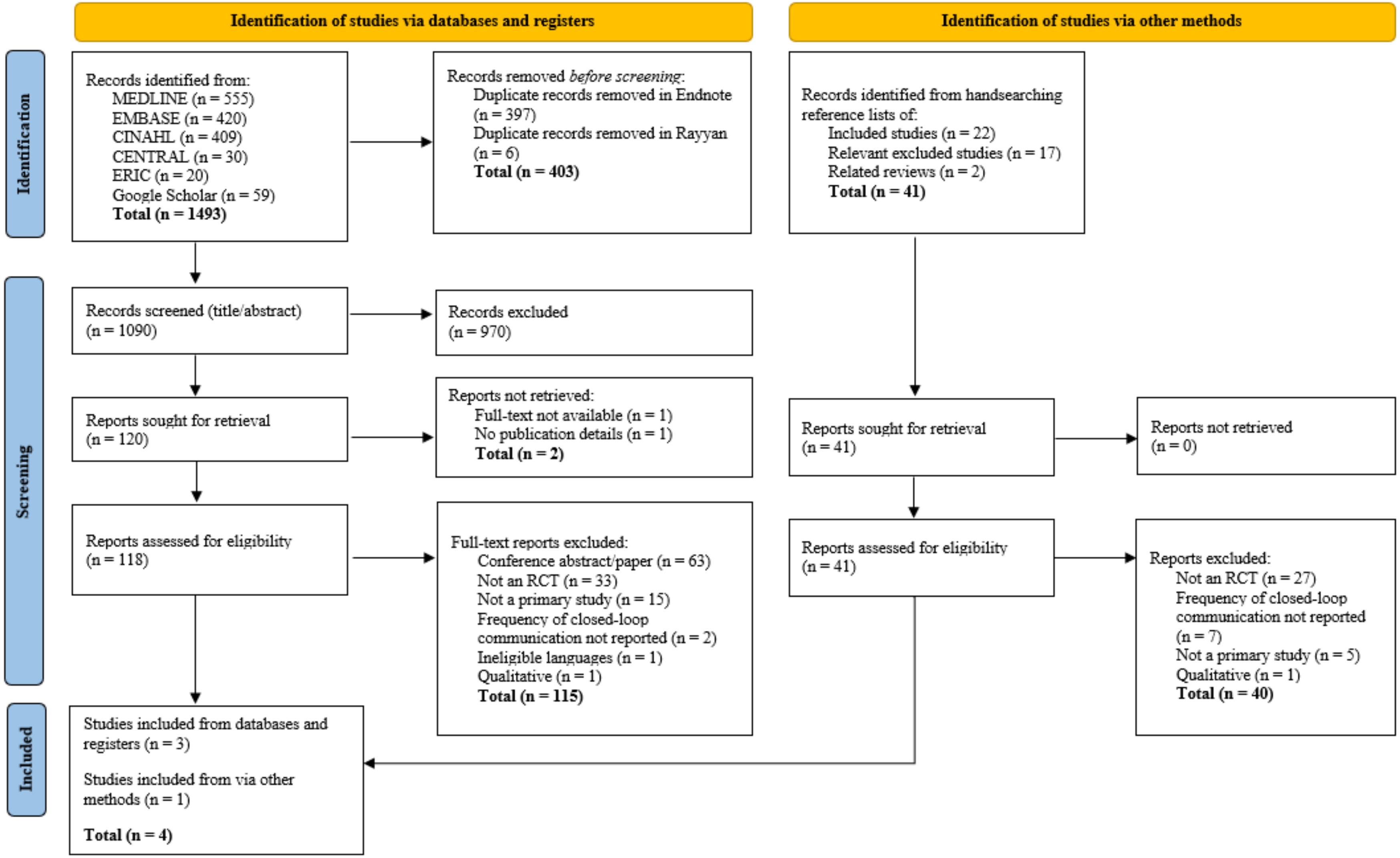
A “Summary of Findings” table was initially planned to be created in a digital program, including the certainty of evidence for the outcome(s), the sum of available data for outcome(s), and the magnitude of effect of the interventions explored (GRADEpro GDT, 2023). However, due to the inclusion of only one outcome in the GRADE assessment and the narrative approach to data synthesis, the table followed the structure exemplified in Murad et al. (2017). Instead of providing explanatory footnotes as planned in the protocol, a detailed description for the judgement of each domain was added in Supplementary appendix 5.

3 Results

3.1 Description of search results

A total of 1493 records were identified through searches in MEDLINE (555), EMBASE (420), CINAHL (409), CENTRAL (30), ERIC (20), and Google Scholar (59). After removing 397 duplicates in Endnote and 6 duplicates in Rayyan, 970 records were excluded by screening titles and abstracts using pre-made screening questions. The remaining 120 records were sought for retrieval, of which 2 records could not be retrieved due to the absence of publication details or not being available in full text. Hence, 118 full text records were read through and assessed according to inclusion/exclusion criteria; 3 studies were included, whereas 115 records were deemed ineligible. An additional 41 records were identified by screening the reference lists of all included studies, related reviews, and relevant excluded studies, and successfully retrieved for full-text reading. Another single study was included, and a further 40 records were excluded. Altogether, there were only four RCTs included in this systematic review, despite the broad inclusion criteria and the comprehensive search. Supplementary appendix 3 presents an overview of exclusion reasons for each individual record assessed in full text.

Figure 1: *PRISMA 2020 flow diagram*



3.2 Description of studies

Table 1 presents a summarization of the most important similarities and differences among the included studies. Supplementary appendix 4 elaborates on these in more detail. All four studies were single-center RCTs and conducted at one single site, available in English, and published between 2019 and 2021. Buyck et al. (2019) was conducted in Switzerland, while the remaining three studies were carried out in the U.S.

3.2.1 Population

Summing up the participants from the studies, there were 197 in total, and the number of participants included in each study varied from 28 to 87. The studies shared a similar recruitment strategy, restricting inclusion to occupational and educational roles either specialized in emergency medicine (EM) or interns focusing on EM as a training component. Furthermore, all studies included physicians, although the composition of participants in terms of education progression varied between studies. Scicchitano et al. (2021) included interns, Hughes et al. (2020) included residents, and both Lopez de Alda et al. (2021) and Buyck et al. (2019) included a combination of residents and fellows. Buyck et al. (2019) also recruited nurses, making this the only one of the four studies to include multiple professions.

3.2.2 Intervention

All studies implemented an intervention of blindfolding the group (team) leader during the replication of acute and critical scenarios, and arranged the training under simulated conditions. Even though the concept of blindfolded simulation training appeared similar across the studies, and all scenarios used could be linked to emergency medicine, the studies' interventions were heterogeneous in multiple aspects. For instance, Lopez de Alda et al. (2021) used a pediatric trauma scenario, while Hughes et al. (2020), Scicchitano et al. (2021), and Buyck et al. (2019) focused on various resuscitation scenarios on either pediatric or adult patient groups, or both. Also, the content of the scenarios differed considerably between the studies.

Another major difference was seen in the intervention set-up. Lopez de Alda et al. (2021), Hughes et al. (2020), and Scicchitano et al. (2021) had all the included participants function as group leader during no more than one or two simulations. The first couple of studies solved the forming of groups by providing three additional embedded participants for each randomized participant. Scicchitano et al. (2021), on the other hand, formed groups before randomization, and within these groups the participants rotated into the leadership role. Buyck et al. (2019) also formed groups prior to randomization, but only one participant in each group acted as group leader during all five simulations carried out in this study.

The studies did not necessarily include the exact same components before, during, and after the simulations, nor did they conduct them in the same order or in a similar manner. The number of simulations and debriefings undergone per participant varied from one to five and one to three, respectively, across the studies. Lopez de Alda et al. (2021) and Hughes et al. (2020) held debriefings post-simulation(s), whereas Buyck et al. (2019) held debriefings after each of the three middle simulations. Scicchitano et al. (2021) did not include debriefing as a training element.

A disparity was also detected across studies in relation to the pre-simulation preparation aspect of the training. Nearly all studies included some sort of presentation or/and instruction on CLC or team dynamics prior to simulation(s), but Buyck et al. (2019) did not specify any specific preparations. Only Hughes et al. (2020) oriented participants to the manikins and simulation lab beforehand. Lopez de Alda et al. (2021), on the other hand, was alone in demonstrating a sample blindfolded simulation to all participants on video prior to the simulation.

The intervention lasted for 4 hours in Buyck et al. (2019), including all parts of the training. Lopez de Alda et al. (2021) and Hughes et al. (2020) only report the duration for certain elements of the training, such as for each individual simulation specifically. No information about the duration of any part of the training has been provided in Scicchitano et al. (2021). In addition, none of the studies that included pre-simulation preparations inform us about how long this aspect of the training took.

3.2.3 Comparison

All studies had no blindfolding as a comparison group and were therefore homogeneous in that regard. Apart from the non-blindfolding element, participants in the comparison group underwent the exact same training as the intervention group.

3.2.4 Outcomes

All studies did a between-group statistical comparison for the outcomes about to be mentioned, and none of the studies did a within-group comparison. As pre-determined in the inclusion criteria, all four studies assessed the frequency of closed-loop communication. Only two of the four studies provided the criteria used to measure a completely executed CLC, and the number of communication steps needed for a loop to be considered fully closed varied (Lopez de Alda et al., 2021; Buyck et al., 2019). Hughes et al. (2020) and Scicchitano et al. (2021) reported the comparison between the blindfolded and non-blindfolded group by a mean value and confidence interval for both groups. Scicchitano et al. (2021) also provided a p-value. The remaining studies reported different statistical metrics, hence not comparable across studies.

Similarly, task performance was also measured by all studies. Two studies measured self-reported performance and Crisis Resource Management (CRM) performance (Lopez de Alda et al., 2021; Hughes et al., 2020). Both studies used the NASA Task Load Index (NASA-TLX) to measure self-reported performance. Yet only one study reported statistical data for the comparison between groups, including a mean value and standard error for each of the groups, and a p-value (Hughes et al., 2020). In regard to CRM performance, Hughes et al. (2020) used the Ottawa Global Rating Scale (GRS) as measurement method to assess leadership skills, problem solving, situational awareness, resource utilization, communication, and overall performance. It is also plausible that Lopez de Alda et al. (2021) used the same scale based on the rationale given in Supplementary appendix 4. None of the studies reported statistical data for the comparison of CRM performance.

Two studies measured time, in seconds, from cardiac arrest to either initiation of chest compression or cardio-pulmonary resuscitation (CPR) (Scicchitano et al., 2021; Buyck et al., 2019), but used different statistical measures. Buyck et al. (2019) reported a median and interquartile range for both groups, measuring the change in time between the first and the last simulation, and a p-value for the between-group comparison. Scicchitano et al. (2021) did not measure a within-group change, but instead assessed differences in the number of seconds used during one simulation by reporting a mean value and confidence interval for both groups, together with a p-value.

Some task performance outcomes were only assessed by individual studies, and not across multiple studies. These include the following: completion of critical actions, both time from cardiac arrest to defibrillation and to first dose of epinephrine, number of pertinent reassessments, and progression of the resuscitation team leader evaluation score assessing various non-clinical skills (Lopez de Alda et al., 2021; Scicchitano et al., 2021; Buyck et al., 2019).

Not all studies assessed medical errors. Communication errors were the only type of error assessed by no more than two studies (Lopez de Alda et al., 2021; Buyck et al., 2019). Lopez de Alda et al., (2021) compared the likelihood of failing the third and fourth step of their criteria for a complete closed-loop communication, by reporting an odds ratios and confidence intervals for each step. Buyck et al. (2019) reported a median and interquartile range for both groups, measuring the change in the number of incomplete communication loops between the first and the last simulation, and a p-value for the between-group comparison.

Table 1: Brief summary of the included studies (N = 4)

Author, year, country	Study design, setting	Population	Intervention, scenario, duration	Comparison	Outcomes
Lopez de Alda et al., 2021 U.S.	RCT Simulated conditions	Total N = 28 Intervention N = 15 Comparison N = 13 EM residents and pediatric EM fellows with Advanced Trauma Life support (ATLS) certification	All participants in this group led one simulation blindfolded (together with three embedded participant nurses) + completed one debriefing session **SS:** pediatric trauma **Preparations:** CLC PowerPoint, video demonstrating a blindfolded example **Duration:** per/simulation (10 minutes), per/debrief (10 minutes)	No blindfolding	Frequency of CLC completion (FoCLC) Missed CLC steps (ME) Completion of critical actions (TP) Self-reported task performance (TP) CRM-performance (TP)
Hughes et al., 2020 U.S.	RCT Simulated conditions	Total N = 34 Intervention N = 17 Comparison N = 17 EM and EM/pediatric dual resident physicians	All participants in this group led two simulations (together with three embedded standardized participants), both blindfolded + completed a one-on-one debriefing	No blindfolding	Frequency of perfect CLC use (FoCLC) Self-reported task performance (TP) CRM-performance (TP)

			SS1: adult resuscitation **SS2:** pediatric resuscitation **Preparations:** video demonstrating CLC, instructions on CLC, orientation to manikin and simulation lab **Duration:** per/simulation (8 minutes) per/debrief (NI)		
Scicchitano et al., 2021 U.S.	RCT Simulated conditions	Total N = 87 Intervention N = 46 Comparison N = 41 Medical interns undergoing advanced cardiovascular life support (ACLS) training	All participants in this group led two simulations (rotating into the leadership role in turns); the practice session, blindfolded, and the testing session, non-blindfolded **SS practice:** resuscitation **SS testing:** resuscitation **Preparations**: presentation about resuscitation team dynamics, standardized instructions **Duration:** per/simulation (NI)	No blindfolding	Number of complete closed-loop communication (FoCLC) Time from cardiac arrest to initiation of chest compression (CPR) (TP) Time from cardiac arrest to defibrillation (TP) Time from cardiac arrest to first dose of epinephrine (TP)

Buyck et al., 2019 Switzerland	RCT Simulated conditions	Total N = 48 Intervention N = 24 Comparison N = 24 Pediatric emergency fellows, pediatric emergency residents, pediatric emergency nurses	All participants partook in five simulations: A, B, C, D, E Each of the simulations, B, C, and D, (1) included the group leaders being blindfolded and (2) were followed by a debrief **SS1-SS5**: pediatric resuscitation **Preparations:** NI **Duration**: intervention (4 hours, including three 20 minutes debriefs),	No blindfolding	Number of complete communication loops (FoCLC) Number of incomplete communication loops (ME) Time to cardio-pulmonary resuscitation from cardiac arrest (TP) Number of pertinent reassessments (TP) Progression of the Resuscitation team leader evaluation score (TP)
N = Number, EM = Emergency medicine, TP = Task performance, ME = Medical error, FoCLC = Frequency of closed-loop communication, SS = Simulation scenario, NI = No information					

3.3 Risk of bias in included studies

All four studies assessed objective outcomes (Lopez de Alda et al., 2021; Hughes et al., 2020; Scicchitano et al., 2021; Buyck et al., 2019). Two studies assessed self-reported task performance (Lopez det Alda et al., 2021; Hughes et al., 2020). A cross-tabulation and graph summarizing all risk of bias judgements was created in RevMan twice; once for objective outcomes, and once for self-reported task performance (Review Manager Web, 2020). These are presented separately in **Figure 2** and **Figure 3**. Support for judgments can be read in Supplementary appendix 4.

3.3.1 Selection bias

All studies mentioned the use of randomization, but not all studies described the randomization method in sufficient detail. Lopez de Alda et al. (2021) used a random number draw, which is considered an adequate method of randomization, and was therefore considered a to have low risk of selection bias from the randomization process. The remaining studies did not fully describe the random component of the sequence generation (Hughes et al., 2020; Scicchitano et al., 2021; Buyck et al., 2019). Hence, these studies were judged to have an unclear risk of bias.

Three studies did not address allocation concealment (Lopez de Alda et al., 2021; Hughes et al., 2020; Scicchitano et al., 2021). Therefore, the risk of bias in all three studies was considered unclear. Buyck et al. (2029) described allocation concealment to some extent by mentioning the use of sealed envelopes, but gave no indications that the envelopes were opaque or sequentially numbered. These additional safeguards had to be ensured and described in order for the allocation concealment to be considered adequate, and thus the risk of bias was also unclear for this specific study.

3.3.2 Performance bias

The lack of blinding of participants and personnel in two of the four studies resulted in a high risk of bias judgement (Lopez de Alda et al., 2021; Scicchitano et al., 2021). Lopez de Alda et al. (2021) gave participants too many clues about the upcoming simulations for a complete blinding to be achieved, while Scicchitano et al. (2021) did not blind the instructors of practice sessions. As all outcomes included in this systematic review measured were behavioral-based, they were particularly susceptible to being influenced by participants' or personnel knowledge or group assignments. Therefore, the lack of blinding in Lopez de Alda et al. (2021) and Scicchitano et al. (2021) is likely to have introduced high risk of performance bias.

In Buyck et al. (2019) the simulation instructor holding debriefing sessions was not blinded, but measures were taken to ensure that this inconvenience did not influence performance. Still, it remains unclear whether participants and instructors involved in the actual simulations were blinded. The same uncertainty goes for Hughes et al. (2020). Hence, both studies were set to have unclear risk of performance bias.

3.3.3 Detection bias

3.3.3.1 Self-reported task performance

Due to this outcome being assessed by participants themselves, the risk of detection bias was considered high in Lopez de Alda et al. (2021) for which blinding of participants was not adequately achieved. The uncertainty surrounding whether participants were blinded in Hughes et al. (2020) also makes it difficult to know for sure whether this affected measurements, thereby reaching the judgement of unclear risk.

3.3.3.2 Objective outcomes

Two studies were assessed to have a high risk of detection bias in regard to objective outcomes (Lopez de Alda et al., 2021; Hughes et al., 2020). Outcome assessors in both studies would have been able to identify participants group assignments from visual cues. Besides, the measurement methods used by the assessors involved some degree of subjective interpretation and judgement. Blinding of outcome assessment was successfully done in Scicchitano et al. (2021) and Buyck et al. (2019). Even if for one of the studies it is unclear whether those who assessed performance were blinded, the study used measurement methods less prone to subjective interpretation (Scicchitano et al., 2021). Therefore, both studies were judged to have low risk of bias.

3.3.4 Attrition bias

3.3.4.1 Self-reported performance

Neither of the two studies that assessed this particular outcome measure reported the number of participants included in the statistical analysis, nor did they provide any information about loss to follow-up (Lopez de Alda et al., 2021; Hughes et al., 2020). Thus, both received the same judgement, namely unclear risk of attrition bias.

3.3.4.2 Objective outcomes

Scicchitano et al. (2021) lost data for only 2% of all participants due to technical difficulties, rather than their true outcomes, meaning that it was reasonable to assess the risk of attrition bias to be low. The three remaining studies, on the other hand, did not provide sufficient information both concerning the number of participants included in the statistical analysis and loss to follow-up (Lopez de Alda et al., 2021; Hughes et al., 2020; Buyck et al. 2019). Therefore, these were judged to have unclear risk of attrition bias.

3.3.5 Reporting bias

No study protocols were found for any of the included studies. However, both Scicchitano et al. (2021) and Buyck et al. (2019) reported results and outcome data for all outcomes listed in the articles' methods section. Therefore, their risk of reporting bias was considered to be low. In contrast, both Lopez de Alda et al. (2021) and Hughes et al. (2020) failed to report important and expected data for multiple outcomes. Lopez de Alda et al. (2021) also did not report results for the completion of critical actions, although assessed. Both of the two last-mentioned studies reached a high risk of bias judgement.

3.3.6 Other biases

For all studies, no other obvious sources of bias were found that could have influenced the intervention and comparison groups differently, and thus led to skewed results.

Figure 2: *RoB 1.0 summary for self-reported performance (left) and objective outcomes (right)*

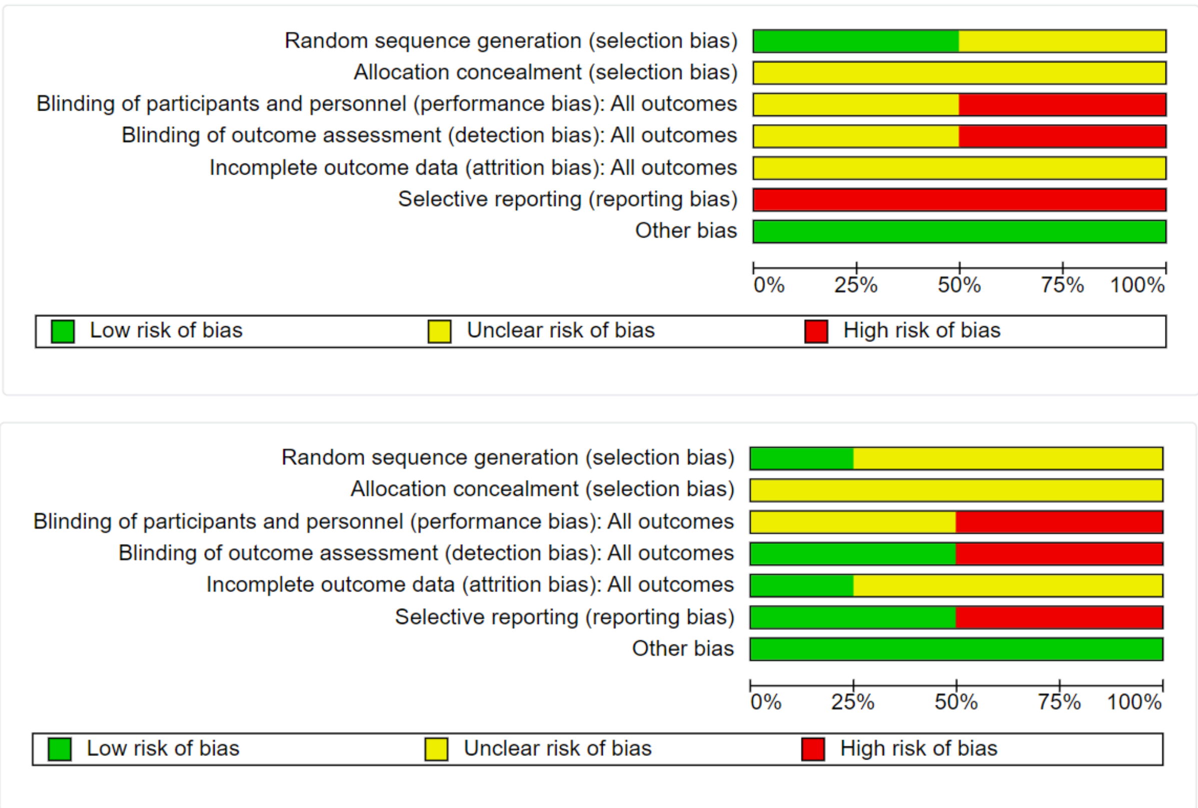


Figure 3: *RoB 1.0 graph for self-reported performance (upper) and objective outcomes (lower)*

3.4 Synthesis of results

As expected, given the broad inclusion criteria, a close inspection of the four studies revealed considerable clinical, methodological, and statistical heterogeneity, and also differences in the statistical metrics used. Based on the details given in Supplementary appendix 4, all studies designed and carried out their blindfolded simulation training differently in terms of including or not including certain elements prior, during, and after the simulation(s). For example, the number of simulations that each participant engaged in also varied. Besides, inconsistent reporting of outcome data and the use of different and statistical metrics for reporting results made it impossible to conduct a meta-analysis. Among the four included studies there were no outcomes measured for which multiple studies reported comparable standardized effect sizes. On these grounds, the alternative narrative approach to synthesize results was adopted instead.

Considering that all included studies had a blindfolded simulation training, but none of them executed the intervention using the exact same methods, no obvious grouping of studies based on their nature of intervention appeared. Yet, an interesting variation was discovered in the tactics used for organizing participants in order to form simulation groups. Two studies formed simulation groups by grouping participants prior to randomization and included only these individuals during the simulation (Scicchitano et al., 2021; Buyck et al., 2019). In contrast, the two other studies did not group participants before randomization, but instead assigned each randomized participant with three embedded participants (Lopez de Alda et al., 2021; Huges et al., 2020). Of presumed interest to those who would like to recreate a blindfolded simulation training in the future, and the possible impact the abovementioned difference may have on the outcome effects, it became reasonable to group studies for synthesis based on this particular setup variation.

The synthesis of results was organized by ordering studies in tables according to their risk of bias summary assessments for either self-reported performance or objective outcomes, depending on which outcome group the individual outcome belonged to. The studies that were judged to have the least overall risk of bias were placed at the top of the tables and mentioned first in the textual synthesis, thereby adhering to a consistent structure. A complete justification for the summarized risk of bias assessments can be found in Supplementary appendix 4. Results were analyzed sequentially by starting with the findings for the primary outcome, then continuing with one secondary outcome measure at a time.

3.4.1 Effects of blindfolded simulation training on the frequency of CLC

Buyck et al. (2019) reported a non-significant difference between the blindfolded group and non-blindfolded group for comparing the within-group change in the number of complete communication loops (p = 0.63). Despite the findings being non-significant, it is important to underscore the tendencies reported as these results stood out from the other studies; the intervention group had no overall median change in the frequency of CLC, whereas the non-blindfolded group in point of fact experienced a change implying increased use.

The three remaining studies reported significant between-group differences (Lopez de Alda et al., 2021; Hughes et al., 2020; Scicchitano et al., 2021). In Lopez de Alda et al. (2021), the odds of closing the loop in the blindfolded group was 13.7 compared to the non-blindfolded group, with a wide though still significant 95% confidence interval ranging from 1.4 to 133.8. Likewise, Scicchitano et al. (2021) found a significant difference in the frequency of CLC favoring the blindfolded group, although this concerned the mean number of complete closed-loop communication (p = 0.002). The mean frequency of closed-loop communication was higher in the blindfolded group versus the non-blindfolded group also for Hughes et al. (2020). Nevertheless, no p-value was reported to confirm the significant difference claimed by the authors, meaning that Lopez de Alda et al. (2021) and Scicchitano et al., (2020) were the only studies to provide sufficient outcome data to verify a significant positive effect of blindfolding on the frequency of CLC. See **Table 2** for detailed descriptions of the results for each individual study.

Table 2*: Between-group differences for the frequency of CLC*

Author, year	**Statistic**	**Intervention group**	**Comparison group**	**P-value**	**RoB 1.0 SA (OO)**
Buyck et al., 2019	Median (IQR)	0 (-7;5)	+3 (-1;4)	0.63.*	Unclear risk
Lopez de Alda et al., 2021	OR (95% CI)	13.7 (1.4-133.8)		NI	High risk
Hughes et al., 2020	Mean value (95% CI)	31.7 (29.34-34.1)	24.6 (21.5-27.7)	NI	High risk
Scicchitano et al., 2021	Mean value (95% CI)	20.3 (18.8-21.8)	16.6 (14.8-18.4)	0.002*	High risk

Odds Ratio = OR, * = p-value <0.005, CI = Confidence interval, NI = No information, RoB 1.0 = Risk of Bias 1.0 tool, SA = Summary assessment, OO = Objective outcomes

3.4.2 Effects of blindfolded simulation training on task performance

3.4.2.1 Self-reported performance

Two studies measured self-reported task performance. Hughes et al. (2020) found a non-significant difference between the blindfolded group and non-blindfolded group when it came to participants' average scores on self-reported performance (p = 0.51). Lopez de Alda et al. (2021) also reported non-significant differences between groups. The last-mentioned study, however, provided no outcome data, nor a p-value for the test of statistical significance verifying the findings. See **Table 3** for detailed descriptions of the results for each individual study.

Table 3*: Between-group differences for self-reported performance*

Author, year	Statistic	Intervention group	Comparison group	P-value	RoB 1.0 SA (SRP)
Hughes et al., 2020	Mean value (SE)	12.29 (1.55)	12.06 (1.16)	0.78	High risk
Lopez de Alda et al., 2021	NI	NI	NI	NI	High risk

SE = Standard Error, * = p-value <0.005, NI = No information, RoB 1.0 = Risk of Bias 1.0 tool, SA = Summary assessment, SRP = Self-reported performance

3.4.2.2 CRM performance

The same two studies that measured self-reported performance also measured CRM performance (Lopez de Alda et al., 2021; Hughes et al., 2020). Both studies found a non-significant difference between the blindfolded group and non-blindfolded group. However, neither study reported outcome data, or a p-value for the test of statistical significance verifying the findings. See **Table 4** for detailed descriptions of the results for each individual study.

Table 4*: Between-group differences for CRM performance*

Author, year	Statistic	Intervention group	Comparison group	P-value	RoB1 SA (OO)
Lopez de Alda et al., 2021	NI	NI	NI	NI	High risk
Hughes et al., 2020	NI	NI	NI	NI	High risk

N/A = No information, RoB 1.0 = Risk of Bias 1.0 tool, SA = Summary assessment, OO = Objective outcomes, * = p-value <0.005

3.4.2.3 Time from cardiac arrest to initiation of chest compression/CPR

Buyck et al. (2019) measured the time from cardiac arrest to CPR, while Scicchitano et al. (2021) measured the time from cardiac arrest to initiation of chest compression. Both studies had a statistically non-significant difference between the blindfolded group and non-blindfolded group for this outcome, as shown from the p-values in Table 5.

Table 5*: Between-group differences for time from cardiac arrest to initiation of chest compression/CPR*

Author, year	Statistic	Intervention group	Comparison group	P-value	RoB 1.0 SA (OO)
Buyck et al., 2019	Median (IQR)	190 (58;267)	61 (17;151)	0.15	Unclear risk
Scicchitano et al., 2021	Mean value (95% CI)	13.96 (10.2-17.8)	13.8 (8.7-19)	0.51	High risk

IQR = Interquartile Range, * = p-value <0.005, CI = Confidence interval, RoB 1.0 = Risk of Bias 1.0 tool, SA = Summary assessment, OO = Objective outcomes

3.4.2.4 Other task performance measures

Several relevant outcomes were assessed by only one out of four studies. Buyck et al. (2019) measured the progression of the resuscitation team leader evaluation score, in regard to non-clinical skills, and did a between-group comparison that appeared to favor the intervention, meaning that the improvement in leadership skills were significantly greater for the blindfolding group (p = 0.04). The same study also reported the effect of blindfolding on the number of pertinent reassessments, but the difference between groups was non-significant (p = 0.57). Scicchitano et al. (2021) measured time from cardiac arrest to both defibrillation (p =

0.51) and first dose of epinephrine (p = 0.82), separately, for which none were statistically significant. Completion of critical actions was assessed by Lopez de Alda et al. (2021), yet no results were reported. Supplementary appendix 4 contains relevant outcome data for all the abovementioned outcomes.

3.4.3 Effects of blindfolded simulation training on medical errors

3.4.3.1 Communication errors

The only type of medical error assessed across the included studies was communication errors. A between-group comparison of incomplete communication loops in Buck et al. (2019) revealed that the blindfolded group had a significant decrease compared to the non-blindfolded group (p = 0.05). Also, in Lopez de Alda et al. (2021) findings were significant, with the non-blindfolded group more likely to fail completing the third and fourth step of the what the authors considered a completely closed communication loop. See **Table 6** for detailed descriptions of the results for each individual study.

Table 6*: Between-group differences for communication errors*

Author, year	**Statistic**	**Intervention group**	**Comparison group**	**P-value**	**RoB 1.0 SA (OO)**
Buyck et al., 2019	Median (IQR)	-2 (-4;-1)	0 (-2;0)	0.05*	Unclear risk
Lopez de Alda et al., 2021	Odds ratio (95% CI)	Step 3 6.2 (1.2-32.0) Step 4 14.6 (2.2-97.6)		NI	High risk

IQR = Interquartile Range, * = p-value <0.005, CI = Confidence interval, RoB 1.0 = Risk of Bias 1.0 tool, SA = Summary assessment, OO = Objective outcomes, NI = No information

3.5 Certainty of evidence

As presented in **Table 7**, the level of certainty in evidence for the frequency of CLC was assessed using the GRADE method and ultimately determined to be very low. Although the two largest studies had the least risk of bias, across all four studies too many domains were assessed as having either high or unclear risk of bias to mitigate concerns about methodological limitations. Moreover, the total number of participants were 197, hence not reaching the minimum 400 threshold (Murad et al., 2017). This, combined with a considerably wide 95% confidence interval reported by Lopez de Alda et al. (2021), raised concerns regarding imprecision. The third downgrade was due to inconsistency both in the magnitude of the effect and because of the considerable heterogeneity between studies in terms of participant characteristics, intervention component, and outcome measurement methods and statistical metrics used. Due to these variations, findings should be considered critically. See Supplementary appendix 5 for detailed descriptions of the judgements made for each GRADE domain.

***Table** 7: Summary of findings (GRADE) for Frequency of CLC*

Outcome	**Effect**	**Number of participants (studies)**	**Certainty in the evidence**
Frequency of CLC	Three studies reported a statistically significant increase and between-group difference in the frequency of CLC. One study showed no median change in the use of CLC in the intervention group but also reported a non-significant difference between groups (Fig. Table 2)	197 (4 RCTs)	Very low ⊕◯◯◯ (Downgraded three times due to risk of bias, imprecision, and inconsistency)

4 Discussion

Several previous systematic reviews have investigated the effects of communication trainings or interventions for healthcare providers on outcomes such as culture, patient safety, self-efficacy, and performance indicators (Alsabri et al., 2022; Lippke et al., 2021; Ardakani et al., 2019). However, according to current understanding, this is the first systematic review to focus specifically on CLC training and consider its effect on CLC frequency, medical errors, and task performance. This presumption is derived from the fact that no systematic review with the exact same approach were discovered during the process of preparing and conducting this thesis, neither through the preliminary searches nor the systematic searches.

4.1 Summary of main results

4.1.1 Communication

Despite the broad inclusion criteria for closed-loop communication training, all four included studies implemented variations of a blindfolded simulation training and compared this up against simulations with no blindfolding. Three out of four RCTs, thus representing a large proportion of the included studies, reported statistically significant findings showing a higher frequency of CLC in the blindfolded group compared to the non-blindfolded group as a result of the intervention. Buyck et al. (2019) and Lopez de Alda et al. (2021) further measured communication errors, with the first study referring to the number of incomplete communication loops and the second study measuring the likelihood of failing certain CLC steps. Both studies found a significant relative decrease in the blindfolded group. These results all infer that such training might encourage effective communicative behavior among healthcare providers providing patient care. Still, one of the included studies reported a larger CLC improvement in the comparison group relative to the intervention group (Buyck et al., 2019). Although this particular between-group difference was non-significant, this indicates an inconsistency in the findings. Furthermore, the non-significant difference essentially means that a conclusion could not be drawn due to lack of adequate data and does not necessarily indicate the absence of a true effect (Dankel, 2019). Even so, this contrasting finding is important not to go unnoticed.

4.1.2 Task performance

A wide range of approaches were used by the included studies to measure task performance. Still, only one study achieved significant results for one singular measure; in Buyck et al. (2019) the intervention group had a comparatively greater improvement in the resuscitation team leader evaluation score. Both studies that reported results for self-reported performance and CRM performance found non-significant differences for these outcome measures. Other findings not reaching statistical significance included time (in seconds) from cardiac arrest to each of the following actions: initiation of chest compression, CPR, defibrillation, and first dose of epinephrine. Should it be that the non-significant results for time-to-task completion were due to there being no true effect, this could have several explanations. The tasks measured in Buyck et al. (2019) and Scicchitano et al. (2021) are typical standard procedures performed as a part of cardiac arrest treatment and may therefore also have been familiar to the participants before the upcoming simulation(s) (Evans et al., 2021). Even if in some cases during the simulation(s) participants were unsure whether they heard the orders correctly, or did not fully understand the orders, their previous experiences with treating cardiac arrests could have enabled them to complete the tasks effectively. This means that, depending on the type of task and the participant executing the task, the absence of CLC will perhaps not always result in the tasks taking longer to complete, even though the likelihood of making errors might increase. Perhaps the effect of CLC training and CLC utilization on time-to-task completion will be greater during simulations for which participants are exposed to unfamiliar tasks or at least unexpected and complex verbal orders.

4.1.3 Risk of bias and certainty of evidence

All studies had high risk of bias for at least one domain, except for Buyck et al. (2019) who had either low or unclear risk of bias for all domains. Therefore, for the grouping of results, none of the studies were judged to have an overall low risk of bias. Besides, the grading of findings linked to the frequency of closed-loop communication also ended up being rated as very low, meaning that the true effect might be substantially different.

4.2 Limitations of the evidence included in this review

4.2.1 Population

Three studies included participants from the U.S., and one study included participants from Switzerland. The geographic scope of the included studies was therefore quite limited.
All participants had either specialization in emergency medicine (EM) or were interns undergoing EM training. Medical teams from emergency medicine or trauma care appear to be the platform frequently chosen by researchers to explore the use of CLC among healthcare providers or test innovative methods for quantifying CLC (Härgestam et al., 2013; Schwindenhammer et al., 2023; Bhangu et al., 2022). The unique high-pressure environment in which these operate may be the driver behind these decisions, as reliance on default communication strategies becomes especially important to ensure optimal patient outcomes in unpredictable and fast-paced situations (Broadfoot & Guth, 2019). Yet, a question remains to whether the results are generalizable to individuals and teams in other specialties. Teams operating in low-paced environments, exposed to fewer unpredictable and urgent situations, and less pressure, might perform differently.

Although all four studies included physicians, the level of education and expertise of the participants, and the extent to which researchers involved participants from multiple professions, varied both within and between studies. This variation could stem from differences between studies in regard to resource availability, practical motives, or based on the researchers' evaluation of which healthcare providers they believe could benefit most from a blindfolded simulation training. Regardless of the reason, evidence from Moeini et al. (2019) suggests that healthcare providers' communication skills vary depending on their educational level. Therefore, it is important to point out the risk that some studies in this systematic review may have started out with better communication skills at baseline than other studies, potentially influencing the degree of improvement in the measured outcomes. More similarity in participant characteristics could presumably have led to greater consistency in the findings.

Bearing in mind that Scicchitano et al. (2021) included only interns, whereas Buyck et al. (2019) included fellows, residents, and nurses, the group compositions also differed between the included studies. As clarified in Supplementary appendix 4, the first-mentioned study recruited participants from interns already undergoing advanced cardiovascular life support training, hence resulting in a rather uniform selection of participants. Scicchitano et al. (2021) mentions this as a limitation for not reflecting the composition of various professions often required during resuscitation. Buyck et al. (2019), on the other hand, aimed for more mixed groups, thus more accurately replicating multi-professional and multidisciplinary teams typically encountered in a clinical setting. An impractical consequence of the different setup methods is that the findings reported by studies investigating the effectiveness of the blindfolded simulation training using homogeneous groups may not be transferable to contexts where the groups are more heterogeneous, and vice versa.

4.2.2 Intervention

Although the inclusion criteria for closed-loop communication training opened up for a variety of approaches, all the included studies investigated the effectiveness of a blindfolded simulation training. This signals that authors of RCTs continue to have faith in this particular intervention method and consider it useful to provide more research evidence. Another advantage is that it enabled a more comprehensive evaluation of the effectiveness of on single intervention concept, despite the small number of studies. One less beneficial aspect of not including multiple types of interventions is the lack of opportunities for comparisons against the blindfolded simulation training. Other relevant closed-loop communication training interventions have indeed been implemented and tested in studies that were not included in this systematic review due to their ineligible study design (Doorey et al., 2022; Diaz & Kimberly, 2020; Emani et al., 2018; Ulmer et al., 2022). Therefore, in order to provide greater insight and meaningful comparisons across different types of interventions, researchers of future systematic reviews should consider expanding inclusion criteria to comprise more study designs.

The nature of the blindfolded simulation training also raises some concerns about the generalizability of the findings. To begin with, this intervention was built on simulation-based training. The rationale for not choosing a real clinical setting undoubtably lies in the concern that blindfolding the team leader or any practitioner during an actual emergency medical event could potentially lead to serious consequences for the patient involved, and therefore constitutes an unethically high risk. Besides, all studies were dependent on replicating a planned number of unwanted medical conditions like trauma or resuscitation, leaving the researchers with simulated scenarios as the obvious solution. Nevertheless, it involves several disadvantages, for instance in regard to the inability of completely mimicking a realistic event (Krishnan et al., 2017; Buyck et al., 2019; Lamé & Dixon-Woods, 2020). During a medical simulation participants may be conscious both of the upcoming event(s) and that no humans' lives are truly at stake, thus influencing their attitude and approach to the simulation scenario (Krishnan et al., 2017). Moreover, the simulations may fail to incorporate crucial steps typically encountered during an actual medical event and cause the communication between healthcare providers to not accurately represent reality (Krishnan et al., 2017).

The controlled environment of simulations can also be considered a limitation. Repeated interruptions, cognitive strain, being responsible for several patients, and decision-making based on incoherent patient information are examples of external influences often present in emergency medical settings, which can have a noticeable influence on both communication and performance. (Broadfoot & Guth, 2019; Lamé & Dixon-Woods, 2020). As these factors are not taken into account during simulated trainings, the behavioral trends observed during simulations cannot necessarily be guaranteed outside such a controlled context.

Generalizability of the evidence may also appear uncertain given that all the included studies performed their blindfolding simulation training at one single site. Multi-center trials, on the other hand, generate effect estimates based on larger and more diverse population across different sites and institutions (Cheng et al., 2018; Bellomo et al., 2009). Therefore, incorporating such studies would have brought more robust evidence to the table. Yet, considering the small number of studies included, as well as their recent publication date, it implies that blindfolded simulation training as an interventional concept were quite recently introduced to clinical research. Larger productions of single-center trials are typically seen in the early phases of testing a new intervention due to its reduced costs and relative simplicity (Bellomo et al., 2009). Perhaps future multi-center trials will use the studies included in this systematic review as a fundament for planning their research.

Despite the fact that all the included studies had a blindfolding training simulation training, the intervention setups differed in multiple aspects. For instance, the simulation scenarios used were highly specific and varied between studies, likely because the researchers wanted to arrange simulations of relevance to their specific participant group. Although all the scenarios could be linked to emergency medicine, the pediatric resuscitation scenarios, and the adult resuscitation scenarios, according to normal procedures, must have proceeded somewhat differently (Hansen et al., 2023). Based on the simulation details revealed in Lopez de Alda et al. (2021) and Buyck et al. (2020), the resuscitation and trauma scenarios did not resemble each other much either. The descriptions further imply that the type of interaction, communicative input, and the activities performed by the participants largely depended on the clinical event that was artificially replicated. Hence, the applicability of the findings may be rather restricted to the specific scenario used, and a question is raised as to whether the studies' reported effects will remain constant regardless of whether a pediatric trauma scenario is replaced with a cardiac catheterization or sepsis event.

The studies' interventions also differed in regard to the structuring of components, number of simulations executed, the organizing of participants, and the incorporation of debriefings and pre-simulation preparations. The multitude of differences underlines both the complexity of the blindfolding simulation training, because multiple training components are included, and the non-standardized and tailored approaches used. Contemplating that more standardized interventions could potentially have led to more consistent and comparable results across the four studies, this heterogeneity constitutes yet another limitation of the evidence.

The usefulness of the evidence in this systematic review is also held back by the lack of reported intervention details. First of all, the length of the entire blindfolding simulation training was reported in only one study (Buyck et al., 2019). Secondly, of the studies that included debriefing as an intervention component, none explained the actual content or structure in sufficient detail to allow for replication. According to a recently published research, inadequate descriptions appear to be a common challenge among studies examining debriefing interventions (Duff et al., 2024). Supposing that others may not be able to repeat the interventions correctly due to incomplete descriptions of the debriefing sections, or are hesitant because the duration is unknown, the findings reported by several of the studies included becomes less valuable.

4.2.3 Comparison

A concern not addressed by any of the four included studies is the potential presence of a Hawthorne effect, referring to a change in behavior due to the participants' awareness of being observed (Berkhout et al., 2022; McCambridge et al., 2014). Considering the nature of the intervention, participants in all the included studies likely expected their simulated scenarios to be both observed and assessed. If the participants in the comparison groups sharpened their communication and performance, despite not being blindfolded, the intention to serve as a control fails. Moreover, the differences between the intervention and comparison groups may have been minimized. This limitation is particularly relevant for this systematic review, as all outcomes considered are behavior-based.

4.2.4 Outcomes

First of all, the certainty of evidence for the frequency of CLC, ended up being very low due to imprecise and inconsistent results, and a considerable risk of bias across the four included studies. Other limitations regarding the results presented for the primary outcome includes the use of difference statistical measures, the lack of reported outcome data to confirm statistical significance, and the variability in the criteria used to measure a complete CLC, although not all studies elaborated on these criteria. Therefore, despite the initially supposed straightforward process of measuring CLC objectively, it appears that the measurement method for doing this has not yet been standardized. Given this information, the slight indication based on the overall findings, that the direction of effect might lean towards greater improvement in the blindfolding group, should be taken with a grain of salt.

These challenges further bring into question whether other techniques for measuring the use of CLC may be more suitable to provide robust and meaningful evidence. More robust and meaningful techniques for measuring the frequency of CLC could therefore perhaps be provided if the reporting were more transparent. Some before-and-after studies have had their participants answer a questionnaire of how often they use closed-loop communication or asked them to rate the CLC ability of their own team (Diaz & Dawson, 2020; Emani et al., 2018). Although more subjective in nature, such measurement methods may be easier to standardize for repeated use across studies. Moreover, by making sure that all participants have an equal understanding of CLC, the subjective influence on the results can be somewhat reduced.

Task performance and medical errors were also important secondary outcomes as the whole point of improving communication is the anticipation that it ultimately has a positive impact on patient care delivery. Communication errors were the only type of medical errors addressed. Hence, this systematic review provides limited evidence of the effect of closed-loop communication training on patient safety indicators more directly linked to patient outcomes, such as errors concerning medication administration or treatment procedures (Alhur et al., 2024; Aghigi et al., 2023).

Two studies found that the blindfolded simulation training significantly reduced communication errors (Lopez de Alda et al., 2021; Buyck et al., 2019). Nevertheless, similar to the frequency of CLC, the results reported by Lopez de Alda et al. (2021) are very imprecise, and hence this evidence should be interpreted with caution. Future research may want to recruit larger sample sizes in order to achieve less imprecise results, as well as more often use comparable statistical measures to enable statistical comparison.

Due to the wide range of approaches used to measure task performance, the included studies seem to disagree about which measurement are most important and relevant. Some studies calculated time-to-task completion, while others assessed CRM performance and self-reported performance. However, common to all studies was that the results for nearly all task-performance measurements were either not reported in sufficient detail or reported as non-significant. Hughes et al. (2020) specifically mentioned that effect sizes for CRM performance were not obtained due to the lack of enough statistical power. According to the researchers' power analysis 53 participants had to be recruited in order to detect an effect, but ended up including no more than 34 participants, thus explaining this shortcoming (Hughes et al. 2020). Buyck et al. (2019) was the only study to measure the progression in the resuscitation team leader evaluation score and found a significant difference between the blindfolding and non-blindfolding group, favoring the intervention. Hence, overall, the systematic review is left with a small amount of evidence that a blindfolded training simulation will improve task performance.

Another shortcoming is the lack of evidence that says something about the long-term effectiveness. None of the included studies had follow-up assessments, a limitation emphasized in each and of the studies (Scicchitano et al., 2021; Buyck et al., 2019; Hughes et al., 2020; Lopez de Alda et al., 2021). Evidence showing sustained changes in CLC frequency, medical errors, and task performance is therefore neither provided nor compares across the studies, even though such understanding is likely to be highly sought after. RCTs are often expensive to conduct and therefore typically lasts for a relatively short period of time (Bosdriesz et al., 2020).

4.3 Limitations of the review process

There are some important limitations about the review process that must be noted. The first concerns the searches performed and the strict inclusion criteria, such as the decisions made to exclude studies published before 1995 and not to search in databases other than MEDLINE, EMBASE, CINAHL, CENTRAL, and ERIC. Besides, due to the short timeframe, unpublished studies, ongoing studies, and grey literature were not searched for. Exclusion further applied to studies not available in English, Norwegian, Danish, or Swedish languages, and those not obtainable in full text. Hence, there is a small possibility that the full range of relevant studies has not been identified and therefore not represented, potentially introducing both selection and publication bias. There is also reason to believe that including only RCTs may have led to a probable loss of valuable evidence and prevented comparison of different types of closed-loop communication training interventions.

If the attempts to obtain additional data from three out of the four included studies had been successful, it is likely that more findings could have been presented in greater detail and included in a more comprehensive meta-analysis. A final limitation important to emphasize, deals with the fact that all parts of the review were completed by only one student, thereby making the prevention of various types of biases through what should ideally be a collaborative work process, difficult. Notwithstanding the foregoing, the careful documentation of each step of the process, and the materials used in that regard, constitutes a key strength in terms of ensuring transparency. The search strategy also took into account that researchers may have used a number of different synonyms and expressions for the term closed-loop communication. Last but not least, it is a strength that the systematic review was conducted and reported following international guidelines.

4.4 Agreements and disagreements with other studies or reviews

Previous systematic reviews have investigated the effectiveness of various types of simulation-based training. However, these simulation-based training interventions do not fulfill the criteria specified in this systematic review for a closed-loop communication training, nor contains any blindfolding element (Ajemba et al., 2024; Fung et al., 2015; Sezgin et al., 2023; Abildgren et al., 2022; Nielsen et al., 2021). Therefore, to the best of my knowledge, no other systematic review has addressed the effectiveness of such blindfolding interventions. The originality of this systematic review contributes important knowledge and insight into emerging closed-loop communication training interventions and thus provides a greater basis for further research. However, it also poses a disadvantage by limiting the opportunities to compare findings and methodological approaches used across similar reviews.

Despite this inconvenience, some of the findings in this systematic review, for instance regarding the frequency of CLC, can nevertheless be compared to relevant before-and-after studies. Diaz & Dawson (2020) and Ulmer et al. (2022) both implemented a blindfolding intervention and measured changes in communicative behavior. Although Diaz & Dawson (2020) used questionnaires to measure the participants' perception of their ability to use CLC, thereby differing slightly from the more objective frequency measurement methods applied among the four included studies, a statistically significant improvement was detected (Diaz & Dawson, 2020). This corresponds to the general findings presented in this systematic review, for which all studies except one reported an increased frequency of CLC in the blindfolded group compared to the non-blindfolded group (Lopez de Alda et al., 2021; Hughes et al., 2020; Scicchitano et al., 2021).

The second-mentioned study on the other hand, found a non-significant difference between the rates of read-back responses observed before and after the blindfolded simulation training

(Ulmer et al., 2022). This refers to the part of the communication loop of which the recipient of a message/order repeats it back, and thus largely captures the essence of CLC (Ulmer et al., 2022). One of the included studies also reported a non-significant result for the frequency of CLC (Buyck et al., 2019). A question therefore remains to whether these two instances of non-significant findings were due to there being no true effect or the lack of enough evidence. Perhaps not all variations of a blindfolding simulation training will produce significant changes in the use of CLC.

A significant reduction in communication errors, although based on a small number of studies, appeared to be more consistent. Lopez de Alda et al. (2021) and Buyck et al. (2019) found that the blindfolded simulation training significantly reduced the likelihood of failing certain CLC steps or the number of incomplete communication loops in the blindfolded group compared to the non-blindfolded group. These findings are supported by the results presented in Ulmer et al. (2022), where an unreliable form of communication was used to a lesser extent following simulation training. Therefore, despite some inconsistency and uncertainty in the evidence, especially with regard the frequency of CLC, several findings suggest a potential positive effect of blindfolded simulation training on participants' overall communication.

Research conducted in an emergency medical context, similarly to the studies included in this review, has shown significant improvements in time-to-task completions for orders using CLC (El-Shafy et al., 2018). Therefore, one can expect that that closed-loop communication training, such as blindfolded simulation training, also increases practitioners' effectiveness indirectly through increased use of CLC. Nonetheless, all results presented in this systematic review on time-to-task completion measurements were non-significant, and thus do not provide any supporting evidence of such an effect. Then again, El-Shafy et al. (2018) measured time to task-completion for medication orders, placement of intravenous lines, administration of intravenous fluids, and laboratory test results. These tasks differ from the cardiac arrest related tasks measured by the studies included in this systematic review, which could possibly explain the varying results.

5 Implications for practice and research

Due to the small number of studies included, the large amount of non-significant, inconsistent, and imprecise results, and the very low certainty of evidence for the primary outcome, the implementation of a blindfolded simulation training should be considered cautiously. The effectiveness might vary substantially depending on the context, considering that all the included studies conducted their training at one single site, and differed in terms of population composition, how the blindfolding simulation training was designed, and which medical scenarios were used during the simulation(s). Replication of the interventions might also prove challenging, as some of the included studies did not describe the debriefing part of their training in sufficient detail.

Future studies should first and foremost ensure that all components of the closed-loop communication training are described in sufficient detail to allow for replication. Researchers are also recommended to thoroughly describe the type of communication exchange considered a CLC, so that similarities and differences in the criteria for CLC can be drawn between studies. There is also a need for more evidence from studies with less risk of bias and from multi-center trials, in order to generate more generalizable findings. Non-significant findings represent a large proportion of the evidence provided in this systematic review. By recruiting larger samples sizes, future studies can ensure that statistically significant differences are detected when present, and provide more precise effect estimates, thereby contributing more meaningful evidence. A more standardized blindfolded simulation training and measurement method for measuring the frequency of CLC will also be needed so that future studies are more comparable.

Evidence concerning long-term effectiveness and retention is presumably meaningful to policy makers, researchers, or leaders from medical institutions or educational programs, who are considering implementing blindfolded simulation training. Therefore, future systematic reviews addressing the exact same question, or investigating the use of another communication strategy, should also consider including study designs that more often contain follow-up assessments.

6 Conclusion

This systematic review found very low certainty of evidence for the effectiveness of a closed-loop communication training method comparing blindfolded versus non-blindfolded simulation training, on the frequency of CLC among healthcare providers. Although most of the findings give the impression of an increased frequency of CLC in the blindfolded groups, thus showing a potential that should not be overlooked, further research is needed to confirm the beneficial impact. Regarding changes in medical errors, the small body of evidence suggests that blindfolded simulation training might lead to a relatively greater reduction in communication errors compared to non-blindfolded simulations. There was a lack of statistically significant improvements in task performance measures, apart from some indication of a positive effect on non-clinical skills among team leaders. Due to several limitations of the evidence provided, the overall findings are not convincing enough to conclude that implementing a blindfolded simulation training as closed-loop communication training for healthcare providers will guarantee increased use of CLC, a reduction in the number medical errors, or improved task performance, or produce the same results as shown in this systematic review, when replicated. Therefore, to best inform about the effectiveness of closed-loop communication training, more research and larger studies are needed that also ideally measure long-term effectiveness and retention.

7 References

Abildgren, L., Lebahn-Hadidi, M., Mogensen, C. B., Toft, P., Nielsen, A. B., Frandsen, T. F., Steffensen, S. V. & Hounsgaard, L. (2022). The effectiveness of improving healthcare teams' human factor skills using simulation-based training: a systematic review. *Advances in Simulation, 7*(12), 1-18. https://doi.org/10.1186/s41077-022-00207-2

Agency for Healthcare Research and Quality. (2023). *TeamSTEPPS Pocket Guide.* Accessed August 2023. https://www.ahrq.gov/sites/default/files/wysiwyg/teamstepps-program/teamstepps-pocket-guide.pdf

Aghighi, N., Aryankhesal, A., Raeissi, P., & Najafpour, Z. (2023). Frequency and influential factors on occurrence of medical errors: A three-year cross-sectional study. *Journal of Education and Health Promotion*, *12*(1), 422–422. https://doi.org/10.4103/jehp.jehp_1726_22

Ajemba, M. N., Ikwe, C., & Iroanya, J., C. (2024). Effectiveness of simulation-based training in medical education: Assessing the impact of simulation-based training on clinical skills acquisition and retention: A systematic review. *World Journal of Advanced Research and Reviews*, *21*(1), 1833–1843. https://doi.org/10.30574/wjarr.2024.21.1.0163

Alhur, A., Alhur, A. A., Al-Rowais, D., Asiri, S., Muslim, H., Alotaibi, D., Al-Rowais, B., Alotaibi, F., Al-Hussayein, S., Alamri, A., Faya, B., Rashoud, W., Alshahrani, R., Alsumait, N., & Alqhtani, H. (2024). Enhancing Patient Safety Through Effective

Interprofessional Communication: A Focus on Medication Error Prevention. *Curēus (Palo Alto, CA)*, *16*(4), e57991–e57991. https://doi.org/10.7759/cureus.57991

Alsabri, M., Boudi, Z., Lauque, D., Dias, R. D., Whelan, J. S., Östlundh, L., Alinier, G., Onyeji, C., Michel, P., Liu, S. W., Jr Camargo, C. A., Lindner, T., Slagman, A., Bates, D. W., Tazarourte, K., Singer, S. J., Toussi, A., Grossman, S., & Bellou, A. (2022). Impact of Teamwork and Communication Training Interventions on Safety Culture and Patient Safety in Emergency Departments: A Systematic Review. *Journal of Patient Safety*, *18*(1), e351–e361. https://doi.org/10.1097/PTS.0000000000000782

Ardakani, M. F., Sharifabad, M. A. M., Bahrami, M. A., & Abargouei, A. S. (2019). The effect of communication skills training on the self-efficacy of nurses: a systematic review and meta-analysis study. *Bali Medical Journal*, *8*(1), 144. https://doi.org/10.15562/bmj.v8i1.1315

Bellomo, R., Warrillow, S. J., & Reade, M. C. (2009). Why we should be wary of single-center trials. *Critical Care Medicine*, *37*(12), 3114–3119. https://doi.org/10.1097/CCM.0b013e3181bc7bd5

Berkhout, C., Berbra, O., Favre, J., Collins, C., Calafiore, M., Peremans, L., & Van Royen, P. (2022). Defining and evaluating the Hawthorne effect in primary care, a systematic review and meta-analysis. *Frontiers in Medicine*, *9*, 1033486–1033486. https://doi.org/10.3389/fmed.2022.1033486

Bhangu, A., Notario, L., Pinto, R. L., Pannell, D., Thomas-Boaz, W., Freedman, C., Tien, H., Nathens, A. B., & da Luz, L. (2022). Closed loop communication in the trauma bay: identifying opportunities for team performance improvement through a video review analysis. *Canadian Journal of Emergency Medicine*, *24*(4), 419–425. https://doi.org/10.1007/s43678-022-00295-z

Bosdriesz, J. R., Stel, V. S., van Diepen, M., Meuleman, Y., Dekker, F. W., Zoccali, C., & Jager, K. J. (2020). Evidence-based medicine—When observational studies are better than randomized controlled trials. *Nephrology (Carlton, Vic.)*, *25*(10), 737–743. https://doi.org/10.1111/nep.13742

Broadfoot, K. J. & Guth, T. A. (2019). Keys to Effective Communication in All Circumstances. In M. E. Moreira & A. French (Ed.), *Communication in Emergency Medicine* (p. 27-42). Oxford University Press.

Burke, C. S., Salas, E., Wilson-Donnelly, K., & Priest, H. (2004). How to turn a team of experts into an expert medical team: guidance from the aviation and military communities. *Quality & Safety in Health Care*, *13*(suppl 1), i96–i104. https://doi.org/10.1136/qshc.2004.009829

Buyck, M., Manzano, S., Haddad, K., Moncousin, A.-C., Galetto-Lacour, A., Blondon, K., & Karam, O. (2019). Effects of Blindfold on Leadership in Pediatric Resuscitation Simulation: A Randomized Trial. *Frontiers in Pediatrics*, *7*, 10–10. https://doi.org/10.3389/fped.2019.00010

Broman, W. C., & Motowidlo, S. J. (1997). Task Performance and Contextual Performance: The Meaning for Personnel Selection Research. *Human Performance*, *10*(2), 99–109. https://doi.org/10.1207/s15327043hup1002_3

Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shepperd, S., Thomas, J., Welch, V., & Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ (Online)*, *368*, l6890–l6890. https://doi.org/10.1136/bmj.l6890

Centre for Reviews and Dissemination. (n.d.). *About PROSPERO.* University of York. https://www.crd.york.ac.uk/prospero/#aboutpage

Cheng, A., Kessler, D., Mackinnon, R., Chang, T. P., Nadkarni, V. M., Hunt, E. A., Duval-Arnould, J., Lin, Y., Pusic, M., & Auerbach, M. (2017). Conducting multicenter research in healthcare simulation: Lessons learned from the INSPIRE network. *Advances in Simulation (London)*, *2*(1), 6–6. https://doi.org/10.1186/s41077-017-0039-0

Crocker, T. F., Lam, N., Jordão, M., Brundle, C., Prescott, M., Forster, A., Ensor, J., Gladman, J., & Clegg, A. (2023). Risk-of-bias assessment using Cochrane's revised tool for randomized trials (RoB 2) was useful but challenging and resource-intensive: observations from a systematic review. *Journal of Clinical Epidemiology*, *161*, 39–45. https://doi.org/10.1016/j.jclinepi.2023.06.015

Dankel, S. J. (2019). What information is provided from non-significant findings and how can this be improved? *Journal of Trainology*, *8*(2), 19–23. https://doi.org/10.17338/trainology.8.2_19

Diaz, M. C. G., & Dawson, K. (2020). Impact of Simulation-Based Closed-Loop Communication Training on Medical Errors in a Pediatric Emergency Department. *American Journal of Medical Quality*, *35*(6), 1-5.

Doorey, A. J., Turi, Z. G., Lazzara, E. H., Casey, M., Kolm, P., Garratt, K. N., & Weintraub, W. S. (2022). Safety gaps in medical team communication: Closing the loop on quality improvement efforts in the cardiac catheterization lab. *Catheterization and Cardiovascular Interventions*, *99*(7), 1953–1962. https://doi.org/10.1002/ccd.30189

Duff, J. P., Morse, K. J., Seelandt, J., Gross, I. T., Lydston, M., Sargeant, J., Dieckmann, P., Allen, J. A., Rudolph, J. W., & Kolbe, M. (2024). Debriefing Methods for Simulation in Healthcare: A Systematic Review. *Simulation in healthcare: journal of the Society for Simulation in Healthcare*, *19*(1S), S112–S121. https://doi.org/10.1097/SIH.0000000000000765

El-Shafy, I. A., Delgado, J., Akerman, M., Bullaro, F., Christopherson, N. A. M., & Prince, J. M. (2018). Closed-Loop Communication Improves Task Completion in Pediatric Trauma Resuscitation. *Journal of Surgical Education*, *75*(1), 58–64. https://doi.org/10.1016/j.jsurg.2017.06.025

Emani, S., Allan, C., Forster, T., Fisk, A., Lagrasta, C., Zheleva, B., Weinstock, P., & Thiagarajan, R. (2018). Simulation training improves team dynamics and performance in a low-resource cardiac intensive care unit. *Annals of Pediatric Cardiology*, *11*(2), 130–136. https://doi.org/10.4103/apc.APC_117_17

Evans, E., Swanson, M. B., Mohr, N., Boulos, N., Vaughan-Sarrazin, M., Chan, P. S., Girotra, S., Grossestreuer, A., Moskowitz, A., Edelson, D., Ornato, J., Peberdy, M. A., Churpek, M., Kurz, M., Anderson Starks, M., Perman, S., & Goldberger, Z. (2021). Epinephrine before defibrillation in patients with shockable in-hospital cardiac arrest: propensity matched analysis. *BMJ (Online)*, *375*, e066534–e066534. https://doi.org/10.1136/bmj-2021-066534

Foronda, C., MacWilliams, B., & McArthur, E. (2016). Interprofessional communication in healthcare: An integrative review. *Nurse Education in Practice*, *19*, 36–40. https://doi.org/10.1016/j.nepr.2016.04.005

Fung, L., Boet, S., Bould, M. D., Qosa, H., Perrier, L., Tricco, A., Tavares, W., & Reeves, S. (2015). Impact of crisis resource management simulation-based training for interprofessional and interdisciplinary teams: A systematic review. *Journal of Interprofessional Care*, *29*(5), 433–444. https://doi.org/10.3109/13561820.2015.1017555

GRADEpro GDT. (2023). *GRADEpro Guideline Development Tool.* [Computer Software]. Available from: https://www.gradepro.org/

Grober, E. D., & Bohnen, J. M. A. (2005). Defining medical error. *Canadian Journal of Surgery, 48*(1), 39-44.

Hansen, M., Walker-Stevenson, G., Bahr, N., Harrod, T., Meckler, G., Eriksson, C., & Guise, J.-M. (2023). Comparison of Resuscitation Quality in Simulated Pediatric and Adult Out-of-Hospital Cardiac Arrest. *JAMA Network Open*, *6*(5), e2313969. https://doi.org/10.1001/jamanetworkopen.2023.13969

Härgestam, M., Lindkvist, M., Brulin, C., Jacobsson, M., & Hultin, M. (2013). Communication in interdisciplinary teams: exploring closed-loop communication during in situ trauma team training. *BMJ Open*, *3*(10), e003525–e003525. https://doi.org/10.1136/bmjopen-2013-003525

Higgins, J, P. T., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions* (Version 5.1.0) [updated March 2011]. The Cochrane Collaboration. Available from: https://handbook-5-1.cochrane.org/

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2023). *Cochrane handbook for systematic reviews of interventions* (Version 6.4) [Updated August 2023]. Cochrane. Available from: www.training.cochrane.org/handbook

Hughes, K. E., Hughes, P. G., Cahir, T., Plitt, J., Ng, V., Bedrick, E., & Ahmed, R. A. (2020). Advanced closed-loop communication training: the blindfolded resuscitation. *BMJ Simulation & Technology Enhanced Learning*, *6*(4), 235–238. https://doi.org/10.1136/bmjstel-2019-000498

Krishnan, D., Keloth, A.V., & Ubedulla, S. (2017). Pros and cons of simulation in medical education: A review. *International Journal of Medical and Health Research, 3*, 84-87.

Lamé, G., & Dixon-Woods, M. (2020). Using clinical simulation to study how to improve quality and safety in healthcare. *BMJ Simulation & Technology Enhanced Learning*, *6*(2), 87–94. https://doi.org/10.1136/bmjstel-2018-000370

Lippke, S., Derksen, C., Keller, F. M., Kötting, L., Schmiedhofer, M., & Welp, A. (2021). Effectiveness of Communication Interventions in Obstetrics-A Systematic Review. *International Journal of Environmental Research and Public Health*, *18*(5), 2616. https://doi.org/10.3390/ijerph18052616

Lopez de Alda, J. X., Patel, N., McNinch, N., & Ahmed, R. A. (2021). A Blindfolded Pediatric Trauma Simulation and Its Effect on Communication and Crisis Resource Management Skills. *Curēus (Palo Alto, CA)*, *13*(11), e19484–e19484. https://doi.org/10.7759/cureus.19484

Makary, M. A., & Daniel, M. (2016). Medical error—the third leading cause of death in the US. *BMJ (Online)*, *353*, i2139–i2139. https://doi.org/10.1136/bmj.i2139

McCambridge, J., Witton, J., & Elbourne, D. R. (2014). Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. *Journal of Clinical Epidemiology*, *67*(3), 267–277. https://doi.org/10.1016/j.jclinepi.2013.08.015

Minozzi, S., Dwan, K., Borrelli, F., & Filippini, G. (2022a). Reliability of the revised Cochrane risk-of-bias tool for randomised trials (RoB2) improved with the use of implementation instruction. *Journal of Clinical Epidemiology*, *141*, 99–105. https://doi.org/10.1016/j.jclinepi.2021.09.021

Minozzi, S., Gonzalez-Lorenzo, M., Cinquini, M., Berardinelli, D., Cagnazzo, C., Ciardullo, S., De Nardi, P., Gammone, M., Iovino, P., Lando, A., Rissone, M., Simeone, G., Stracuzzi, M., Venezia, G., Moja, L., Costantino, G., Cianciulli, A., Cinnirella, A., Grosso, F., … Zambarbieri, E. (2022b). Adherence of systematic reviews to Cochrane RoB2 guidance was frequently poor: a meta epidemiological study. *Journal of Clinical Epidemiology*, *152*, 47–55. https://doi.org/10.1016/j.jclinepi.2022.09.003

Moeini, B., Abasi, H., Otogara, M. & Akbarzadeh, M. (2019). Communication Skills and Its Related Factors Among Medical Staff. *Hormogozan medical journal. 23* (1), e86254.

Murad, M. H., Mustafa, R. A., Schünemann, H. J., Sultan, S., & Santesso, N. (2017). Rating the certainty in evidence in the absence of a single estimate of effect. *Evidence-Based Medicine (English Ed.)*, *22*(3), 85–87. https://doi.org/10.1136/ebmed-2017-110668

Nielsen, R. P., Nikolajsen, L., Paltved, C., & Aagaard, R. (2021). Effect of simulation-based team training in airway management: a systematic review. *Anaesthesia*, *76*(10), 1404–1415. https://doi.org/10.1111/anae.15375

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210–210. https://doi.org/10.1186/s13643-016-0384-4

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLoS Medicine*, *18*(3), e1003583–e1003583. https://doi.org/10.1371/journal.pmed.1003583

Peyre, S.-E. (2014). CRICO Operating Room Team Training Collaborative: Closed Loop Communication. https://www.rmf.harvard.edu/Risk-Prevention-and-

Education/Article-Catalog-Page/Articles/2014/CRICO-Operating-Room-Team-Training-Collaborative-Closed-Loop-Communication

Review Manager Web. (2020). *RevMan Web* (Version 1.22.1) [Computer Software]. The Cochrane Collaboration. Available from: https://training.cochrane.org/online-learning/core-software/revman

Schünemann, H., Brozek, J., Guyatt, G. & Oxman, A. (2013). Grade Handbook for grading quality of evidence and strength of recommendations. The GRADE Working Group. https://gdt.gradepro.org/app/handbook/handbook.html

Schwindenhammer, V., Rimmelé, T., Duclos, A., Haesebaert, J., Lilot, M., & Abraham, P. (2023). A new standardized tool for quantification of closed-loop communication in trauma care: CAST Grid reliability study. *Injury*, *54*(9), 110851–110851. https://doi.org/10.1016/j.injury.2023.110851

Scicchitano, E., Stark, P., Koetter, P., Michalak, N., & Zurca, A. D. (2021). Blindfolding Improves Communication in Inexperienced Residents Undergoing ACLS Training. *Journal of Graduate Medical Education*, *13*(1), 123–127. https://doi.org/10.4300/JGME-D-20-00620.1

Sezgin, M. G., & Bektas, H. (2023). Effectiveness of interprofessional simulation-based education programs to improve teamwork and communication for students in the healthcare profession: A systematic review and meta-analysis of randomized controlled trials. *Nurse Education Today*, *120*, 105619–105619. https://doi.org/10.1016/j.nedt.2022.105619

Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., … Higgins, J. P. T. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ (Online)*, *366*, l4898–l4898. https://doi.org/10.1136/bmj.l4898

The Endnote Team. (2013). *Endnote (Version 21).* [Computer Software] Clarivate. https://endnote.com/product-details?language=en

Ulmer, F. F., Lutz, A. M., Müller, F., Riva, T., Bütikofer, L., & Greif, R. (2022). Communication Patterns During Routine Patient Care in a Pediatric Intensive Care Unit: The Behavioral Impact of In Situ Simulation. *Journal of Patient Safety*, *18*(2), e573–e579. https://doi.org/10.1097/PTS.0000000000000872

World Health Organization. (2021). *Global Patient Safety Action Plan 2021-2030: Toward elimination avoidable harm in health care.* https://www.who.int/publications/i/item/9789240032705

Zimmer, M., Czarniecki, D. M., & Sahm, S. (2021). Communication of preclinical emergency teams in critical situations: A nationwide study. *PloS One*, *16*(5), e0250932–e0250932. https://doi.org/10.1371/journal.pone.0250932

8 Supplementary material

8.1 Supplementary appendix 1: Full search strategy

MEDLINE

Database(s): Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations, Daily and Versions (1946 to January 10, 2023)
Date searched: January 11, 2024
Search type: Advanced
Total hits: 555

#	Searches	Results
1	"closed-loop communication".ab,kf,ti.	143
2	"feedback*".ab,kf,ti.	20
3	"checkback*".ab,kf,ti.	1
4	"callout*".ab,kf,ti.	99
5	"call-out*".ab,kf.ti.	395
6	(read* adj3 back*).ab,kf,ti.	2389
7	(check* adj3 back*).ab,kf,ti.	2285
8	(repeat* adj3 back*).ab,kf,ti.	1977
9	(clos* adj3 loop*).ab,kf,ti.	18481
10	2 or 3 or 4 or 5 or 6 or 7 or 8 or 9	25595
11	exp communication/ or exp teach-back communication/	371027
12	exp interpersonal relations/ or interprofessional relations/ or exp interdisciplinary communication/	353218
13	11 or 12	630512
14	10 and 13	500
15	1 or 14	581
16	Limit 15 to yr="1995 -Current"	555

Embase

Database(s): Embase Classic+Embase (1946 to 2024 January 10)
Date searched: January 11, 2024
Search type: Advanced
Total hits: 420

#	Searches	Results
1	"closed-loop communication".ab,kf,ti.	284
2	"readback*".ab,kf,ti.	29
3	"checkback*".ab,kf,ti.	3
4	"callout*".ab,kf,ti.	216
5	"call-out*".ab,kf.ti.	635

6	(read* adj3 back*).ab,kf,ti.	3589
7	(check* adj3 back*).ab,kf,ti.	5446
8	(repeat* adj3 back*).ab,kf,ti.	2902
9	(clos* adj3 loop*).ab,kf,ti.	23989
10	2 or 3 or 4 or 5 or 6 or 7 or 8 or 9	36718
11	*interpersonal communication/ or exp communication skill/	73478
12	exp interdisciplinary communication	13753
13	11 or 12	86958
14	10 and 13	195
15	1 or 14	425
16	Limit 15 to yr=”1995 -Current”	420

CINAHL

Database(s): EBSCOhost CINAHL
Date searched: January 14, 2024
Search type: Advanced
Total hits: 409

Search ID#	Search terms	Search Options	Actions
S17	S1 OR S15	**Limiters** - Publication Date: 19950101-20241231 **Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (409)
S16	S1 OR S15	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (418)
S15	S10 AND S14	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (382)
S14	S11 OR S12 OR S13	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (586,325)
S13	(MM "Intraprofessional Relations")	**Expanders** - Apply equivalent subjects	View results (5,118)

		Search modes - Boolean/Phrase	
S12	(MH "Interpersonal Relations+") OR (MH "Interprofessional Relations+")	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (326,751)
S11	(MH "Communication+") OR (MM "Communication Skills Training") OR (MM "Communication Skills")	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (332,802)
S10	S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (3,290)
S9	TI (clos* n3 loop*) OR AB (clos* n3 loop*)	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (2,443)
S8	TI (repeat* w3 back*) OR AB (repeat* w3 back*)	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (127)
S7	TI (check* w3 back*) OR AB (check* w3 back*)	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (97)
S6	TI (read* w3 back*) OR AB (read* w3 back*)	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (303)
S5	TI "call-out*" OR AB "call-out*"	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (267)
S4	TI "callout*" OR AB "callout*"	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (65)

S3	TI "checkback*" OR AB "checkback*"	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (1)
S2	TI "readback*" OR AB "readback*"	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (8)
S1	TI "closed-loop communication" OR AB "closed-loop communication"	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (76)

ERIC

Database(s): EBSCOhost ERIC
Date searched: January 10, 2024
Search type: Advanced
Total hits: 20

Search ID#	Search terms	Search Options	Actions
S14	S1 OR S12	**Limiters** - Publication Date: 19950101-20241231 **Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (20)
S13	S1 OR S12	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (25)
S12	S10 AND S11	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (24)
S11	DE "Interpersonal Communication" OR DE "Communication Skills" OR DE "Communication (Thought Transfer)" OR DE "Interpersonal Communication" OR DE "Communication Strategies"	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (46,357)

S10	S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (1,066)
S9	TI (clos* n3 loop*) OR AB (clos* n3 loop*)	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (300)
S8	TI (repeat* w3 back*) OR AB (repeat* w3 back*)	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (23)
S7	TI (check* w3 back*) OR AB (check* w3 back*)	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (41)
S6	TI (read* w3 back*) OR AB (read* w3 back*)	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (632)
S5	TI "call-out*" OR AB "call-out*"	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (62)
S4	TI "callout*" OR AB "callout*"	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (8)
S3	TI "checkback*" OR AB "checkback*"	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (0)
S2	TI "readback*” OR AB “readback*”	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (2)
S1	TI "closed-loop communication" OR AB "closed-loop communication"	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	View results (2)

CENTRAL

Database(s): Cochrane Central Register of Controlled Trials
Date searched: January 10, 2024
Search type: Advanced
Total hits: 30 (1 Cochrane Review and 29 Trials)

#	Search manager	Results
1	(closed-loop NEXT communication):ti,ab,kw (Word variations have been searched)	**17**
2	(read-back*):ti,ab,kw (Word variations have been searched)	**6**
3	(checkback*):ti,ab,kw (Word variations have been searched)	**2**
4	(callout*):ti,ab,kw (Word variations have been searched)	**27**
5	(call-out*):ti,ab,kw (Word variations have been searched)	**112**
6	(read* NEXT/3 back*):ti,ab,kw (Word variations have been searched)	**42**
7	(check* NEXT/3 back*):ti,ab,kw (Word variations have been searched)	**33**
8	(repeat* NEXT/3 back*):ti,ab,kw (Word variations have been searched)	**63**
9	(clos* NEAR/2 loop*):ti,ab,kw (Word variations have been searched)	**1909**
10	#2 or #3 or #4 or #5 or #6 or #7 or #8 or #9	**2164**
11	MeSH descriptor: [Communication] explode all trees	**12783**
12	#10 and #11	**17**
13	#12 or #1 with Cochrane Library publication date from Jan 1995 to Dec 2024	**30**

Google Scholar

Date searched: January 10, 2024
Search type: Advanced
Total hits: 59

Return articles dated between: 1995-2024
Find articles with the exact phrase: closed loop communication (allintitle)

8.2 Supplementary appendix 2: Screening questions

For title and abstract screening

#	Questions	Yes	Unclear	No
1	Is it a primary study?			
2	Is the population healthcare providers?			
3	Is it about communication training?			
4	Is it a quantitative study?			

Exclude if the answer to at least one question is “No”
Include if the answers to all questions are “Yes” or “Unclear” (or a mix of both).

For title and abstract screening

#	Questions	Yes	No
	Population		
1	Is the population healthcare providers involved in direct patient care or assistance? *They may be healthcare students/workers, health professionals (also those undergoing postgraduate or specialized training)*		
2	Is the population only involved in the human field of healthcare?		
3	Does at least 80% of the population fit the inclusion criteria above?		
	Intervention		
4	Is the training designed to improve effective communication, with emphasis on verification to ensure that messages/orders are understood correctly?		
5	**Only answer when relevant:** If the study is about implementing new policies, systems, guidelines, checklists, procedures, protocols, or digital devices, is complementary communication training also a part of the intervention?		
	Comparison		
6	Does the study include one of the following comparison groups: standard training, no training, or another intervention?		
	Outcome		
7	Is frequency of closed-loop communication reported as an outcome?		
	Study design		
8	Is it an RCT?		
	Other criteria		

9	Is it a primary study?		
10	Is it a quantitative study?		
11	Is the study available in English, Norwegian, Swedish, or Danish?		
12	Is the publication year between 1995 and 2024		

Exclude if the answer to at least one question is “No”
Include if the answers to all questions are “Yes”

8.3 Supplementary appendix 3: Reasons for excluded studies read in full text

Total number of excluded studies from database and registry searches (n=115)

#	Excluded studies	Reasons
1	Ahmed R., Hughes K., & Hughes P. (2018). The blindfolded code training exercise. *The Clinical Teacher, 15*(2), 120-125. https://doi.org/10.1111/tct.12639	Not an RCT
2	Alexandrino, H., Baptista, S., Vale, L., Junior, J. H. Z., Espada, P. C Junior, D. S., Vane, L. A., Carvalho, V. H., Marcelo, L., Madeira, F Duarte, R., Ferreira, L., Pereira, J., Pinheiro, L. F., Fraga, G. P., & Mesquita, C. (2020). Improving Intraoperative Communication in T The Educational Effect of the Joint DSTC TM-DATC TM Courses. *Journal of Surgery*, *44*(6), 1856-1862. https://doi.org/https://dx.doi.org/10.1007/s00268-020-05421-5	Not an RCT
3	Anonymous. (2004). Complying with JCAHO's read-back requirement. *Hospital Peer Review*, *29*(6), 78, 83-74.	Not a primary study
4	Arista, M. C., Hella, M., & Robinson, C. (2020). Non-technical factors affecting student's status epilepticus simulation experience in a neurology clerkship [Conference Abstract]. *Neurology. Conference: 72nd Annual Meeting of the American Academy of Neurology, AAN*, *94*(15 Supplement).	Conference abstract/paper
5	Backlund, E.-L. (2014). Kommunikationsmodellers påverkan på kommunikationen och patientsäkerheten i akuta situationer: SBAR och Closed-Loop Communication.	Not a primary study
6	Bank, I., Snell, L., & Bhanji, F. (2014). Pediatric crisis resource management training improves emergency medicine trainees' perceived ability to manage emergencies and ability to identify teamwork errors. *Pediatric Emergency Care*, *30*(12), 879-883. https://doi.org/https://dx.doi.org/10.1097/PEC.0000000000000302	Not an RCT
7	Barsuk, J. H., Paparello, J., & Cotts, W. (2012). Appropriate diuretic dosing: closed loop communication [Editorial Comment]. *Journal of Hospital Medicine (Online)*, *7*(3), 167-169. https://doi.org/https://dx.doi.org/10.1002/jhm.1000	Not a primary study
8	Benefits of Closed-Loop Communication. (2020). *AACN Bold Voices*, *12*(4), 20-20.	Not a primary study
9	Bhatnagar, S., Szerlip, M., Lotun, K., Abidov, A., Subramanian, S., Sethi, G., Paidy, S., & Poston, R. (2013). Improving team performance and patient outcomes during Transcatheter Aortic Valve Replacement (TAVR) through simulation [Conference Abstract]. *Catheterization and Cardiovascular Interventions*, *1)*, S151-S152. https://doi.org/https://dx.doi.org/10.1002/ccd.24919	Conference abstract/paper
10	Bikak, A. L., Turi, Z., Plusch, K., Weintraub, W., Garratt, K., & Doorey, A. (2017). Effective communication in the cardiac catheterization laboratory. Analysis of closed loop communication in a critical care setting [Conference Abstract]. *Catheterization and*	Conference abstract/paper

	Cardiovascular Interventions, *89(Supplement 2)*, S148. https://doi.org/https://dx.doi.org/10.1002/ccd.27053	
11	Blankenship, J. C. (2020). Communication to cure cath chaos [Editorial Comment]. *Catheterization & Cardiovascular Interventions*, *95*(5), E154-E155. https://doi.org/https://dx.doi.org/10.1002/ccd.28687	Not a primary study
12	Bouhabel, S., Nhan, C., Kay-Rivest, E., Nguyen, L. H. P., & Bank, I. (2015). Error detection as a means to assess improvement in a CRM workshop [Conference Abstract]. *Otolaryngology - Head and Neck Surgery (United States)*, *1)*, 48. https://doi.org/https://dx.doi.org/10.1177/0194599815593290c	Conference abstract/paper
13	Boyd, M., Cumin, D., Lombard, B., Torrie, J., Civil, N., & Weller, J. (2014). Read-back improves information transfer in simulated clinical crises [Research Support, Non-U.S. Gov't]. *BMJ Quality & Safety*, *23*(12), 989-993. https://doi.org/https://dx.doi.org/10.1136/bmjqs-2014-003096	Not an RCT
14	Broman, K. K., Kensinger, C., Hart, H., Mathisen, J., & Kripalani, S. (2017). Closing the loop: a process evaluation of inpatient care team communication [Research Support, U.S. Gov't, Non-P.H.S.]. *BMJ Quality & Safety*, *26*(1), 30-32. https://doi.org/https://dx.doi.org/10.1136/bmjqs-2015-004580	Not an RCT
15	Brown, J. P. (2004). Closing the communication loop: using readback/hearback to support patient safety. *Joint Commission Journal on Quality & Safety*, *30*(8), 460-464.	Not a primary study
16	Brown, S. J. (2009). The importance of closing the loop. *Family Practice Management*, *16*(4), 32.	Not a primary study
17	Chan, C. K., So, E. H., Ng, G. W., Ma, T. W., Chan, K. K., & Ho, L. Y. (2016). Participant evaluation of simulation training using crew resource management in a hospital setting in Hong Kong. *Hong Kong Medical Journal*, *22*(2), 131-137. https://doi.org/https://dx.doi.org/10.12809/hkmj154595	Not an RCT
18	Chin, T. L., Cash, J., Blacker, H., Thomas, M., Bernal, N. P., Burton, K., & Joe, V. C. (2019). Timely debriefing facilitates real-time communication and feedback, improves team performance, and provides data clarity for quality improvement [Conference Abstract]. *Journal of Burn Care and Research*, *40(Supplement 1)*, S201. https://doi.org/https://dx.doi.org/10.1093/jbcr/irz013.351	Conference abstract/paper
19	Chladek, M. S., Doughty, C., Patel, B., Alade, K., Rus, M., Shook, J., & K, L. I.-W. (2021). The Standardisation of handoffs in a large academic paediatric emergency department using I-PASS. *BMJ Open Quality*, *10*(3), 07. https://doi.org/https://dx.doi.org/10.1136/bmjoq-2020-001254	Not an RCT
20	Cory, M. J., Hebbar, K. B., Colman, N., Pierson, A., & Clarke, S. A. (2020). Multidisciplinary Simulation-Based Team Training: Knowledge Acquisition and Shifting Perception. *Clinical Simulation in Nursing*, *41*, 14-21. https://doi.org/10.1016/j.ecns.2020.01.001	Not an RCT
21	Couloures, K. G., & Allen, C. (2017). Use of Simulation to Improve Cardiopulmonary Resuscitation Performance and Code	Not an RCT

	Team Communication for Pediatric Residents. *Mededportal Publications*, *13*, 10555. https://doi.org/https://dx.doi.org/10.15766/mep_2374-8265.10555	
22	Dabney, C., Appling, N. A., & Herr, M. J. (2020). An Interprofessional Branching Simulation to Introduce RN First Assistant Students to Their Role in the Perioperative Setting. *AORN journal*, *112*(5), 471-477. https://doi.org/https://dx.doi.org/10.1002/aorn.13211	Not an RCT
23	DeJohn, P. (2009). Closing the loop on biopsy specimens. *OR Manager*, *25*(1), 28-29.	Not a primary study
24	Del Pozo, A., Giustiniano, E., Fusilli, N., Simili, V., Calabro, L., Pugliese, L., Stomeo, N., Ebm, C., & Brusa, S. (2021). COVID-19 emergencies management learning through high-fidelity simulation for medical residents-Humanitas Simulation Centre Experience [Conference Abstract]. *Intensive Care Medicine Experimental. Conference: European Society of Intensive Care Medicine Annual Congress, ESICM*, *9*(SUPPL 1). https://doi.org/https://dx.doi.org/10.1186/s40635-021-00413-8	Conference abstract/paper
25	Diaz, M. C. G., & Dawson, K. (2020). Impact of Simulation-Based Closed-Loop Communication Training on Medical Errors in a Pediatric Emergency Department. *American Journal of Medical Quality*, *35*(6), 474-478. https://doi.org/https://dx.doi.org/10.1177/1062860620912480	Not an RCT
26	Diaz, M. C. G., & Dawson, K. (2018). Simulation based training to improve closed loop communication in a pediatric emergency department [Conference Abstract]. *Pediatrics. Conference: National Conference on Education*, *144*(2). https://doi.org/https://dx.doi.org/10.1542/peds.144.2-MeetingAbstract.156	Conference abstract/paper
27	Dickinson, M., Hatch, M., Henson, A., Gooderham, L., & Cupitt, J. (2018). Paediatric in-situ simulation on adult intensive therapy unit [Conference Abstract]. *Journal of the Intensive Care Society*, *19(2 Supplement 1)*, 140. https://doi.org/https://dx.doi.org/10.1177/1751143718772957	Conference abstract/paper
28	Dintzis, S., Mehri, S., Struijk, J., Luff, D., Rendi, M., & Raab, S. (2013). Improved pathology resident performance in critical value communication using simulation based communication training [Conference Abstract]. *Laboratory Investigation*, *1)*, 126A. https://doi.org/https://dx.doi.org/10.1038/labinvest.2013.18	Conference abstract/paper
29	Dixon, R., Gonzales, J., Shah, N., Davenport, J., Wilson, T., Kapoor, N., Yeung, S. Y. A., Heavner, M., & Tisherman, S. (2019). Interprofessional simulation for enhancing knowledge, teamwork, and communication in students [Conference Abstract]. *Critical Care Medicine. Conference: 48th Critical Care Congress of the Society of Critical Care Medicine, SCCM*, *47*(1 Supplement 1)	Conference abstract/paper
30	Doorey, A. J., Turi, Z. G., Lazzara, E. H., Casey, M., Kolm, P., Garratt, K. N., & Weintraub, W. S. (2022). Safety gaps in medical	Not an RCT

	team communication: Closing the loop on quality improvement efforts in the cardiac catheterization lab [Observational Study]. *Catheterization & Cardiovascular Interventions*, *99*(7), 1953-1962. https://doi.org/https://dx.doi.org/10.1002/ccd.30189	
31	Doorey, A. J., Turi, Z. G., Lazzara, E. H., Mendoza, E. G., Garratt, K. N., & Weintraub, W. S. (2020). Safety gaps in medical team communication: Results of quality improvement efforts in a cardiac catheterization laboratory [Observational Study]. *Catheterization & Cardiovascular Interventions*, *95*(1), 136-144. https://doi.org/https://dx.doi.org/10.1002/ccd.28298	Not an RCT
32	Durst, J., Temming, L., Gamboa, C., Tuuli, M., Macones, G., & Young, O. (2017). Improving interprofessional communication utilizing obstetric simulation training [Conference Abstract]. *Obstetrics and Gynecology*, *130(Supplement 1)*, 47S-48S. https://doi.org/https://dx.doi.org/10.1097/01.AOG.0000525745.88585.40	Conference abstract/paper
33	Dutta, A., Moffett, B., Mobeen, S., & Singh, A. (2019). Reducing antibiotic use in children with respiratory syncitial virus-related bronchiolitis: Implementation of TeamSTEPPS 2.0 to improve pharmacy-physician communication in a community hospital antibiotic stewardship program [Conference Abstract]. *Open Forum Infectious Diseases*, *6(Supplement 2)*, S408. https://doi.org/https://dx.doi.org/10.1093/ofid/ofz360.1009	Conference abstract/paper
34	Eisenberg, M., & Garson, G. (2006). Closing the Loop. SPHERE brings EMS & public health together. *Journal of Emergency Medical Services*, *31*(6), 56-59.	Conference abstract/paper
35	El-Shafy, I. A., Delgado, J., Akerman, M., Bullaro, F., Christopherson, N. A. M., & Prince, J. M. (2018). Closed-Loop Communication Improves Task Completion in Pediatric Trauma Resuscitation [Observational Study]. *Journal of Surgical Education*, *75*(1), 58-64. https://doi.org/https://dx.doi.org/10.1016/j.jsurg.2017.06.025	Not an RCT
36	Emani, S., Allan, C., Fisk, A., Lagrasta, C., Zheleva, B., Weinstock, P., & Thiagarajan, R. (2017). Simulation training engenders higher comfort levels among caregivers and lowers time to response in low resource intensive care unit [Conference Abstract]. *Cardiology in the Young*, *27(4)*, S343-S344. https://doi.org/https://dx.doi.org/10.1017/S104795111700110X	Conference abstract/paper
37	Emani, S. S., Allan, C. K., Forster, T., Fisk, A. C., Lagrasta, C., Zheleva, B., Weinstock, P., & Thiagarajan, R. R. (2018). Simulation training improves team dynamics and performance in a low-resource cardiac intensive care unit. *Annals of Pediatric Cardiology*, *11*(2), 130-136. https://doi.org/10.4103/apc.APC_117_17	Not an RCT
38	Espeso, L., Meyenburg, T., Bell, R., & Herr, D. (2016). Nurse-driven implementation of cardiac surgery advanced life support in a cardiac surgery ICU [Conference Abstract]. *Critical Care Medicine*, *44(12 Supplement 1)*, 177.	Conference abstract/paper

	https://doi.org/https://dx.doi.org/10.1097/01.ccm.0000509082.44657.8b	
39	Franzova, E., Waldrop, G., Nwankwo, I., Ader, J., Bell, M., Velazquez, A., Agarwal, S., Roh, D., Park, S., Connolly, S., Claassen, J., Amiel, J., & Ghoshal, S. (2023). Utility and Feasibility of Simulation-based Training in Teaching Concepts of Traumatic Brain Management to Neurology Trainees [Conference Abstract]. *Neurology. Conference: American Academy Of Neurology Annual Meeting, AAN*, *100*(17 Supplement 2). https://doi.org/https://dx.doi.org/10.1212/WNL.0000000000203982	Conference abstract/paper
40	Fulton, C., Saravana-Bawan, B., Riley, B., Widder, S., & Paton-Gay, D. (2017). Techniques in crisis resource management teaching [Conference Abstract]. *CMAJ. Canadian Medical Association Journal*, *60(3 Supplement 2)*, S30. https://doi.org/https://dx.doi.org/10.1503/cjs.005617	Conference abstract/paper
41	Glater-Welt, L., Haddad, D., Esperanza, M., Sweberg, T., Kessel, A., Bakar, A., Nishisaki, A., & Gangadharan, S. (2018). Crm-based "takeoff" script reduces esophageal intubations in a multidisciplinary PICU [Conference Abstract]. *Critical Care Medicine*, *46(Supplement 1)*, 647. https://doi.org/https://dx.doi.org/10.1097/01.ccm.0000529330.17231.35	Conference abstract/paper
42	Greeves, R., & Henderson, J. (2016). Implementing the shift model of handover: Shifting patient care [Conference Abstract]. *Archives of Disease in Childhood*, *101(Supplement 1)*, A323. https://doi.org/https://dx.doi.org/10.1136/archdischild-2016-310863.530	Conference abstract/paper
43	Gurley, K. L., Bagg, N. M., Tibbles, C. D., Greenberg, P., Song, E., Janes, M., Yu, W., & Rosen, C. L. (2016). An innovative quality and safety curriculum for emergency medicine residents [Conference Abstract]. *Academic Emergency Medicine*, *1)*, S198. https://doi.org/https://dx.doi.org/10.1111/acem.12974	Conference abstract/paper
44	Hargestam, M., Lindkvist, M., Brulin, C., Jacobsson, M., & Hultin, M. (2013). Communication in interdisciplinary teams: exploring closed-loop communication during in situ trauma team training. *BMJ Open*, *3*(10), e003525. https://doi.org/https://dx.doi.org/10.1136/bmjopen-2013-003525	Not an RCT
45	Hargestam, M., Lindkvist, M., Jacobsson, M., Brulin, C., & Hultin, M. (2016). Trauma teams and time to early management during in situ trauma team training. *BMJ Open*, *6*(1), e009911. https://doi.org/https://dx.doi.org/10.1136/bmjopen-2015-009911	Not an RCT
46	Hartman, J. A., Anderson, D. M., Ding, J., & Keech, J. C. (2022). Interprofessional veno-veno bypass simulation improved team confidence. *Surgery*, *171*(4), 904-907. https://doi.org/https://dx.doi.org/10.1016/j.surg.2021.08.005	Not an RCT
47	Hassan, M., & Ali, N. (2023). A loop closure Audit on Quality of Handover for Patient care to the Post-Anaesthetic Care Unit. *Irish Medical Journal*, *116*(9), 855	Not an RCT

48	Hazwani, T., Ashraf, N., Hasan, Z., Antar, M., Kazzaz, Y., & Alali, H. (2020). Effect of a pediatric mock code on resuscitation skills and team performance: An in situ simulation experience over three years [Conference Abstract]. *European Journal of Emergency Medicine*, *27(SUPPL 1)*, e15-e16. https://doi.org/https://dx.doi.org/10.1097/MEJ.0000000000000743	Conference abstract/paper
49	Heidari, A., & Khandani, A. K. (2006). Improved closed-loop communication in the presence of feedback delay and error. *IIIE 2006 40th Annual Conference on Information Sciences and Systems,* 209-294.	Conference abstract/paper
50	Hochman, K. A., Adler, N., Volpicelli, F., Wertheimer, B., Zabar, S., & Szyld, D. (2016). Kick-starting a culture of safety: How teamstepps and simulation transformed attitudes on the medicine service [Conference Abstract]. *Journal of General Internal Medicine*, *1)*, S288.	Conference abstract/paper
51	Holt, A., Nahabedian, L., Patel, A., & Copeman, A. (2020). Closing the loop to break the ice: improving communication through fun introductory activities [Conference Abstract]. *BMJ Simulation and Technology Enhanced Learning*, *6(Supplement 1)*, A22.	Conference abstract/paper
52	Huckels-Baumgart, S., Niederberger, M., Manser, T., Meier, C. R., & Meyer-Massetti, C. (2017). A combined intervention to reduce interruptions during medication preparation and double-checking: a pilot-study evaluating the impact of staff training and safety vests. *Journal of Nursing Management*, *25*(7), 539-548. https://doi.org/10.1111/jonm.12491	Not an RCT
53	Hughes, K. E., Hughes, P. G., Cahir, T. M., Plitt, J., Ng, V., Bedrick, E., & Ahmed, R. A. (2019). 249 Crisis Resource Management Training: the Blindfolded Resuscitation [Journal article; Conference proceeding]. *Annals of Emergency Medicine*, *74*(4), S98. https://doi.org/10.1016/j.annemergmed.2019.08.417	Conference abstract/paper
54	Hughes, P. G., Hughes, K. E., & Ahmed, R. A. (2017). I can literally do this blindfolded: The blindfolded code training simulation exercise [Conference Abstract]. *Academic Emergency Medicine*, *24(Supplement 1)*, S297. https://doi.org/https://dx.doi.org/10.1111/acem.13204	Conference abstract/paper
55	Jackson, J. W., Ofsthun, N. J., Meredith, C., Goldberg, M. E., Onta, U., & Maddux, F. W. (2017). Reduction in rates of blood stream infection associated with adoption of teamstepps as a framework for improved hemodialysis facility workflows [Conference Abstract]. *Journal of the American Society of Nephrology*, *Conference: Kidney Week 2017. New Orleans, LA United States*. *28*, 329.	Conference abstract/paper
56	Kabi, A., Dhar, M., Arora, P., Bhardwaj, B. B., Chowdhury, N., & Rao, S. (2021). Effectiveness of a Simulation-Based Training Program in Improving the Preparedness of Health Care Workers Involved in the Airway Management of COVID-19 Patients. *Cureus*, *13*(8), e17323. https://doi.org/https://dx.doi.org/10.7759/cureus.17323	Not an RCT

57	Kavanagh, K., Carter, B., & Azzimonti, B. (2018). Crisis resource management simulation during otolaryngology residency training: Do non-technical skills transfer into early practice? Pilot study [Conference Abstract]. *Otolaryngology - Head and Neck Surgery (United States)*, *159(1 Supplement 1)*, P288. https://doi.org/https://dx.doi.org/10.1177/0194599818787193g	Conference abstract/paper
58	Keane, O. A., Chambers, C., Brady, C. M., Rehberg, J., Iyer, S., & Santore, M. T. (2023). Reducing Retained Foreign Objects in the Operating Room: A Quality Improvement Initiative. *Journal of the American College of Surgeons*, *237*(6), 864-872. https://doi.org/https://dx.doi.org/10.1097/XCS.0000000000000847	Not an RCT
59	Kennedy, J. L., Jones, S. M., Porter, N., White, M. L., Gephardt, G., Hill, T., Cantrell, M., Nick, T. G., Melguizo, M., Smith, C., Boateng, B. A., Perry, T. T., Scurlock, A. M., & Thompson, T. M. (2013). High-fidelity hybrid simulation of allergic emergencies demonstrates improved preparedness for office emergencies in pediatric allergy clinics [Research Support, Non-U.S. Gov't]. *The Journal of Allergy & Clinical Immunology in Practice*, *1*(6), 608-617.e601-614. https://doi.org/https://dx.doi.org/10.1016/j.jaip.2013.07.006	Not an RCT
60	Kennedy, J., Jones, S., Perry, T., Scurlock, A., Smith, C., & Thompson, T. (2010). Community training in pediatric allergy clinics using high fidelity simulation demonstrates improved preparation for office emergencies [Conference Abstract]. *Journal of Investigative Medicine*, *58(2)*, 458. https://doi.org/https://dx.doi.org/10.231/JIM.0b013e3182820c55	Conference abstract/paper
61	Kerr, B., Hoey, R., Gaunt, D., & Holding, M. (2020). Simulated resuscitative thoracotomy training for emergency department middle grades [Conference Abstract]. *BMJ Simulation and Technology Enhanced Learning*, *6(Supplement 1)*, A37-A38. https://doi.org/https://dx.doi.org/10.1136/bmjstel-2020-aspihconf.63	Conference abstract/paper
62	Kessler, C. S., Tadisina, K. K., Saks, M., Franzen, D., Woods, R., Banh, K. V., Bounds, R., Smith, M., Deiorio, N., Schwartz, A., & Deiorio, N. (2015). The 5Cs of Consultation: Training Medical Students to Communicate Effectively in the Emergency Department. *Journal of Emergency Medicine (0736-4679)*, *49*(5), 713-721. https://doi.org/10.1016/j.jemermed.2015.05.012	Frequency of closed-loop communication not reported
63	Lamba, S., Nagurka, R., Offin, M., & Compton, S. (2012). Case-based simulation: Critical conversations around resuscitation of the critically ill or injured patient [Conference Abstract]. *Annals of Emergency Medicine*, *1)*, S149-S150. https://doi.org/https://dx.doi.org/10.1016/j.annemergmed.2012.06.457	Conference abstract/paper
64	Lantz, J., Torzone, A., Allard, J., Storey, J., & Koch, J. (2011). From chaos to confidence: Improving multidisci-plinary staff competence in resuscitation onto ECMO [Conference Abstract]. *Pediatric Critical Care Medicine*, *1)*, S88. https://doi.org/https://dx.doi.org/10.1097/PCC.0b013e318223ad3	Conference abstract/paper

65	Lee, C., Bernstein, P., Chazotte, C., Angert, R., Bernstein, J., McGowan, A., Merkatz, I., & Goffman, D. (2013). Interdisciplinary obstetric simulation training improves team performance [Conference Abstract]. *American Journal of Obstetrics and Gynecology*, *208(1 SUPPL.1)*, S318. https://doi.org/https://dx.doi.org/10.1016/j.ajog.2012.10.095	Conference abstract/paper
66	Lee, D. D., Colwill, A. C., Teel, J., & Srinivas, S. K. (2018). Safe Passage: Improving the Transition of Care Between Triage and Labor and Delivery [Observational Study]. *Quality management in health care*, *27*(4), 223-228. https://doi.org/https://dx.doi.org/10.1097/QMH.0000000000000191	Not an RCT
67	Leonardi, M., Espada, M., & Condous, G. (2021). Closing the communication loop between gynecological surgeons, diagnostic imaging experts and pathologists in endometriosis: building bridges between specialties. *Ultrasound in Obstetrics & Gynecology*, *57*(4), 523-525. https://doi.org/https://dx.doi.org/10.1002/uog.23595	Not a primary study
68	Macri, J., Perrodin, J., Jenson, W., Wright, F., & Asher, Z. (2024). 1088: Multidisciplinary code simulation and closed-loop communication utilization. *Critical Care Medicine*, *52*(1), S516.	Conference abstract/paper
69	Maio Matos, F., Bento, M., Ralha, T., Jesus, J., Paiva, M., & Martins, M. (2014). Simulation instructor training to improve patient safety translation to clinical setting-crisis resource management, debriefing, speak up and close-loop communication [Conference Abstract]. *European Journal of Anaesthesiology*, *52)*, 252.	Conference abstract/paper
70	McCune, T. L. N., Poirier, T. M. P., Buckley, M., & Durkee, R. (2021). Transition of Hospital Based Endoscopy Nurses to an Ambulatory Setting...American Society of PeriAnesthesia Nurses 40th National Conference, 25-29 April, 2021. *Journal of PeriAnesthesia Nursing*, *36*(4), e12-e12. https://doi.org/10.1016/j.jopan.2021.06.040	Conference abstract/paper
71	McDonald, T. M., Wang, B. Y., & Lee, T. K. (2017). Critical communication between pathologists and surgeons during intraoperative consultations: Direct face-to-face interaction [Conference Abstract]. *Laboratory Investigation*, *97(Supplement 1)*, 513A. https://doi.org/https://dx.doi.org/10.1038/labinvest.2016.182	Conference abstract/paper
72	McGowan, L., Riley, J., Piccoli, L., & Patterson, D. (2022). In situ interprofessional pediatric simulation study in the emergency department [Conference Abstract]. *Western Journal of Emergency Medicine*, *23(1.1)*, S57-S58.	Conference abstract/paper
73	Mehri, S., Wingerson, A., Struijk, J., Luff, D., Raab, S., & Dintzis, S. (2013). Effectiveness of pathology resident communication simulation based education on clincial practice [Conference Abstract]. *Laboratory Investigation*, *1)*, 478A. https://doi.org/https://dx.doi.org/10.1038/labinvest.2013.36	Conference abstract/paper

74	Monks, S. C., & Maclennan, K. (2016). Human factor training for the multidisciplinary team [Conference Abstract]. *International Journal of Obstetric Anesthesia*, *1)*, S34.	Conference abstract/paper
75	Moonan, M. (2015). Using simulation to teach transplant nurses about posterior reversible encephalopathy syndrome(PRES) [Conference Abstract]. *Pediatric Transplantation*, *1)*, 137.	Conference abstract/paper
76	Nastars, D., Burns, L., Rangel, A., Schultz, R., & Cates, L. A. (2016). An innovative use of respiratory students to gain experience in high stakes neonatal simulations. *Respiratory Care*, *61*(10), OF25-OF25.	Conference abstract/paper
77	Nelson, A., Vahali, S., Kornegay, J., & Yarris, L. (2016). Teaching EPA 10: A simulated clinical experience improves novice medical student knowledge and comfort in recognizing patients requiring emergent care [Conference Abstract]. *Western Journal of Emergency Medicine*, *17(Supplement 1)*, S32-S33.	Conference abstract/paper
78	O'Brien, M., Hogan, C., Ahern, J., & Almodovar, M. (2009). Communication and team building skills in the cardiovascular intensive care unit [Conference Abstract]. *Cardiology in the Young*, *2)*, 157. https://doi.org/https://dx.doi.org/10.1017/S1047951109991739	Conference abstract/paper
79	Olm-Shipman, C. M., Tan, E., Miller, S., Bennett, A. V., Brissie, M., Sardinha-Soares, S., & Jordan, J. D. (2017). Implementation of team-based communication strategies improves interprofessional collaboration in the neurosciences intensive care unit [Conference Abstract]. *Neurocritical Care*, *27(2 Supplement 1)*, S114-S115. https://doi.org/https://dx.doi.org/10.1007/s12028-017-0465-9	Conference abstract/paper
80	Ortiz Figueroa, F., Moftakhar, Y., Dobbins Iv, A. L., Khan, R., Dasgupta, R., Blanda, R., Marchand, T., & Ahmed, R. (2016). Trauma Boot Camp: A Simulation-Based Pilot Study. *Cureus*, *8*(1), e463. https://doi.org/https://dx.doi.org/10.7759/cureus.463	Not an RCT
81	Peddy, S. B. (2018). Acute Hypoxemia in Infants With Cyanotic Complex Cardiac Anatomy: Simulation Cases for Pediatric Fellows. *Mededportal Publications*, *14*, 10706. https://doi.org/https://dx.doi.org/10.15766/mep_2374-8265.10706	Not an RCT
82	Pena, F., Pereira, V. C., Zamberlan, C., Rangel, R. F., & Ilha, S. (2020). Closed loop communication and blue code in the emergency department: elaboration of a Standard Operating Procedure. *Research, Society and Development*, *9*(4), 5.	Ineligible languages
83	Piasio, N. A., Santos, L., Carr, A., Garza, S., & Muniz, J. (2018). Ipass handoff: decreasing communication failures by utilizing a standardized patient handoff for neonatal and pediatric transport [Conference Abstract]. *Pediatrics. Conference: National Conference on Education*, *144*(2). https://doi.org/https://dx.doi.org/10.1542/peds.144.2_MeetingAbstract.880	Conference abstract/paper
84	Plitt, J., & Ng, V. (2017). The mass casualty incident: A simulation-based curriculum [Conference Abstract]. *Annals of Emergency Medicine*, *70(4 Supplement 1)*, S160.	Conference abstract/paper

85	Pop, S. R., Butala, N., & Heymann, W. R. (2019). Closing the Loop: Assuring That Skin Biopsy Results Are Received by Patients Following Hospital Discharge. *American Journal of Medical Quality*, *34*(1), 96. https://doi.org/https://dx.doi.org/10.1177/1062860618794419	Conference abstract/paper
86	Prabhakar, H., Cooper, J. B., Sabel, A., Weckbach, S., Mehler, P. S., & Stahel, P. F. (2012). Introducing standardized "readbacks" to improve patient safety in surgery: a prospective survey in 92 providers at a public safety-net hospital. *BMC surgery*, *12*, 8. https://doi.org/https://dx.doi.org/10.1186/1471-2482-12-8	Not an RCT
87	Read-backs might help -- do you have the time? (2008). *ED Management*, *20*(9), 99-99. https://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=105671655&site=ehost-live	Not a primary study
88	Repplinger, D., Kearney, T., Olson, K., & Smollin, C. (2013). Implementation and evaluation of high-fidelity simulation cases for the acutely poisoned patient: An inter-professional approach [Conference Abstract]. *Clinical Toxicology*, *51(4)*, 370-371. https://doi.org/https://dx.doi.org/10.3109/15563650.2013.785188	Conference abstract/paper
89	Rosser, A., Sullivan, S., Thompson, R., & Jung, H. S. (2019). Automated natural language processing of closed-loop communication in trauma resuscitations [Conference Abstract]. *Critical Care Medicine. Conference: 48th Critical Care Congress of the Society of Critical Care Medicine, SCCM*, *47*(1 Supplement 1)	Conference abstract/paper
90	Ruebsam, M. L., Metelmann, B., Hofmann, C., Orsson, D., Hahnenkamp, K., & Metelmann, C. (2023). Bilingual resuscitation training does not affect adherence to resuscitation guidelines but reduces leadership skills and overall team performance. An observational study with cross-border German-Polish training. *Resuscitation Plus*, *15*, 100436. https://doi.org/https://dx.doi.org/10.1016/j.resplu.2023.100436	Not an RCT
91	Safety, M., LeVan, J., Gonzaba, G., Solis, S., & Seitz, R. (2018). Utilizing rapid cycle deliberate practice simulation to improve nicu resuscitation [Conference Abstract]. *Pediatrics. Conference: National Conference on Education*, *144*(2). https://doi.org/https://dx.doi.org/10.1542/peds.144.2-MeetingAbstract.161	Conference abstract/paper
92	Schaller, D., Trumph, C., & Foldie, J. (2022). Resuscitation leadership: An introductory curriculum for the 4th year medical student [Conference Abstract]. *Western Journal of Emergency Medicine*, *23(1.1)*, S64.	Conference abstract/paper
93	Schweickert, W. D., Jablonski, J., Bayes, B., Chowdhury, M., Whitman, C., Tian, J., Blette, B., Tran, T., & Halpern, S. D. (2023). Structured Mobilization for Critically Ill Patients: a Pragmatic Cluster-Randomized Trial [Journal article]. *American journal of respiratory and critical care medicine*, *208*(1), 49-58. https://doi.org/10.1164/rccm.202209-1763OC	Frequency of closed-loop communication not reported

94	Sensmeier, J. (2009). Close the loop on med errors. *Nursing Management*, *40*(11 Suppl), 2-3. https://doi.org/https://dx.doi.org/10.1097/01.NUMA.0000366144.77496.8a	Not a primary study
95	Seto, A. H. (2022). Closing the loop in cath lab communication: Avoiding the tower of babble [Comment]. *Catheterization & Cardiovascular Interventions*, *99*(7), 1963-1964. https://doi.org/https://dx.doi.org/10.1002/ccd.30232	Not a primary study
96	Severson, M. A., Maxson, P. M., Wrobleski, D. S., & Dozois, E. J. (2014). Simulation-based team training and debriefing to enhance nursing and physician collaboration [Research Support, Non-U.S. Gov't]. *Journal of continuing education in nursing*, *45*(7), 297-303; quiz 304-295. https://doi.org/https://dx.doi.org/10.3928/00220124-20140620-03	Qualitative
97	Siddiqi, M. F., Vanderbilt, A., Siddiqi, N., Tanner, N. T., Huggins, J. T., Doelken, P., & Pastis, N. (2014). The use of simulation training for teaching airway management to pulmonary/critical care fellows [Conference Abstract]. *American Journal of Respiratory and Critical Care Medicine*. *Conference: American Thoracic Society International Conference, ATS*, *189* (MeetingAbstracts).	Conference abstract/paper
98	Siddiqi, N., Tanner, N. T., Siddiqi, M. F., Vanderbilt, A., & Pastis, N. (2014). Leadership and crew resource management qi project for emergent airway management among pulmonary/critical care fellows [Conference Abstract]. *American Journal of Respiratory and Critical Care Medicine*. *Conference: American Thoracic Society International Conference, ATS*, *189* (MeetingAbstracts).	Conference abstract/paper
99	Simon, A., DeAngelis, M., Schreyer, K., Blome, A., & Healy, M. (2020). Sign out down the alley: A novel workshop-based approach to teaching ED transitions of care [Conference Abstract]. *Western Journal of Emergency Medicine*, *21(4.1)*, S68.	Conference abstract/paper
100	Smith, C. J., Wadman, M. C., Harrison, J., & Beck, G. L. (2015). Evaluation of a brief handoff skills workshop targeted at incoming interns [Conference Abstract]. *Journal of General Internal Medicine*, *2)*, S495-S496.	Conference abstract/paper
101	Srinivas, S. V., Ratwani, A., Davison, K., O'Brien, S. R., & Jamieson, D. B. (2018). Teaching the intangible: Implementation of a simulation program to improve communication and teamwork during code blue responses at an academic medical center [Conference Abstract]. *American Journal of Respiratory and Critical Care Medicine*. *Conference: American Thoracic Society International Conference, ATS*, *197* (MeetingAbstracts).	Conference abstract/paper
102	Thompson, T., Pasala, S., Gephardt, G., Boateng, B., & Schexnayder, S. (2011). Pediatric critical care and emergency medicine regional boot camp improves fellow confidence in essential skills and communication [Conference Abstract]. *Critical Care Medicine*, *12)*, 18. https://doi.org/https://dx.doi.org/10.1097/01.ccm.0000408627.24229.88	Conference abstract/paper

103	Todd, C., Zvonar, I., Palmer, M., Mayo, K., & Mittelman, A. (2023). Calling a Consult Like a Pro: A Novel Workshop to Make Your Interns Shine [Conference Abstract]. *Academic Emergency Medicine*, *30(Supplement 1)*, 404-405. https://doi.org/https://dx.doi.org/10.1111/acem.14718	Conference abstract/paper
104	Tofil, N. M., Morris, J. L., Peterson, D. T., Watts, P., Epps, C., Harrington, K. F., Leon, K., Pierce, C., & White, M. L. (2014). Interprofessional simulation training improves knowledge and teamwork in nursing and medical students during internal medicine clerkship [Research Support, Non-U.S. Gov't]. *Journal of Hospital Medicine (Online)*, *9*(3), 189-192. https://doi.org/https://dx.doi.org/10.1002/jhm.2126	Not an RCT
105	Tsai, P. D. (2023). Closed Loop Communication.	Not a primary study
106	Ulmer, F. F., Lutz, A. M., Muller, F., Riva, T., Butikofer, L., & Greif, R. (2022). Communication Patterns During Routine Patient Care in a Pediatric Intensive Care Unit: The Behavioral Impact of In Situ Simulation [Observational Study]. *Journal of patient safety*, *18*(2), e573-e579. https://doi.org/https://dx.doi.org/10.1097/PTS.0000000000000872	Not an RCT
107	Viggers, S., Andersen, B. R., & Lippert, A. (2015). Introduction of the Immediate Life Support course in Denmark [Conference Abstract]. *Resuscitation*, *1)*, 93. https://doi.org/https://dx.doi.org/10.1016/j.resuscitation.2015.09.221	Conference abstract/paper
108	Ward, B. (2020). Close the loop on diagnostic test results. *Patient Safety Monitor Journal*, *21*(3), 1-3. https://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=141695434&site=ehost-live	Not a primary study
109	Ward, B. (2020). Close the loop on test results. *Medical Environment Update*, *30*(3), 4-5. https://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=141695370&site=ehost-live	Not a primary study
110	Whitman, C., & Bhajandas, N. (2021). Effectiveness of a zoom-based virtual advanced cardiac life support simulation activity [Conference Abstract]. *Critical Care Medicine*, *49(1 SUPPL 1)*, 473. https://doi.org/https://dx.doi.org/10.1097/01.ccm.0000729696.58150.7a	Conference abstract/paper
111	Wiese, M., Assioun, J., DiPalma, S., Sadowski, M., Koehler, A. D., Emborsky, M., Kramer, B., & Wrotniak, B. (2021). Implementation of an Institutional Simulation Curriculum to Teach Team Leadership Skills, Management of Critically Ill Patients, and Emergent Procedures to Pediatric Residents [Conference Abstract]. *Pediatrics. Conference: National Conference and Exhibition Meeting of the American Academy of Pediatrics, AAP*, *149*(no pagination).	Conference abstract/paper
112	Winterbottom, F. A., & Webre, H. (2021). Rapid Response System Restructure: Focus on Prevention and Early Intervention. *Critical*	Not an RCT

	Care Nursing Quarterly, *44*(4), 424-430. https://doi.org/https://dx.doi.org/10.1097/CNQ.0000000000000379	
113	Xu, J., Dong, X., Yin, H., Guan, Z., Li, Z., Qu, F., Chen, T., Wang, C., Fang, Q., & Zhang, L. (2022). Improve Cardiac Emergency Preparedness by Building a Team-Based Cardiopulmonary Resuscitation Educational Plan [Journal article]. *Frontiers in public health*, *10*, 895367. https://doi.org/10.3389/fpubh.2022.895367	Not an RCT
114	Yan, D. H., Anderson, I. M., Bhola, M., Bacevice, A. M., & James, M. Y. (2016). Teaching resuscitation skills to pediatric residents through simulation [Conference Abstract]. *Academic Pediatrics*, *16(6)*, e52.	Conference abstract/paper
115	Zhong, Y., Lyu, G., He, X., Zhang, Y., & Ge, S. S. (2023). Distributed Active Fault-Tolerant Cooperative Control for Multiagent Systems With Communication Delays and External Disturbances. *IEEE transactions on cybernetics*, *53*(7), 4642-4652. https://doi.org/https://dx.doi.org/10.1109/TCYB.2021.3133463	Not an RCT

Total number of excluded studies from handsearching the reference lists of all included studies, related reviews, and relevant excluded studies (n = 40)

#	Excluded studies	Reasons (found)
1	Aggarwal, R., Mytton, O. T., Derbrew, M., Hananel, D., Heydenburg, M., Issenberg, B., MacAulay, C., Mancini, M. E., Morimoto, T., Soper, N., Ziv, A., & Reznick, R. (2010). Training and simulation for patient safety. *Quality & safety in health care*, *19 Suppl 2*, i34–i43. https://doi.org/10.1136/qshc.2009.038562	Not a primary study
2	Allan, C. K., Thiagarajan, R. R., Beke, D., Imprescia, A., Kappus, L. J., Garden, A., Hayes, G., Laussen, P. C., Bacha, E., & Weinstock, P. H. (2010). Simulation-based training delivered directly to the pediatric cardiac intensive care unit engenders preparedness, comfort, and decreased anxiety among multidisciplinary resuscitation teams. *The Journal of thoracic and cardiovascular surgery*, *140*(3), 646–652. https://doi.org/10.1016/j.jtcvs.2010.04.027	Not an RCT
3	Andreatta, P., Saxton, E., Thompson, M., & Annich, G. (2011). Simulation-based mock codes significantly correlate with improved pediatric patient cardiopulmonary arrest survival rates. *Pediatric Critical Care Medicine*, *12*(1), 33–38. https://doi.org/10.1097/PCC.0b013e3181e89270	Not an RCT
4	Awad, S. S., Fagan, S. P., Bellows, C., Albo, D., Green-Rashad, B., De La Garza, M., & Berger, D. H. (2005). Bridging the communication gap in the operating room with medical team training. *The American Journal of Surgery*, *190*(5), 770–774. https://doi.org/10.1016/j.amjsurg.2005.07.018	Not an RCT
5	Blanié, A., Gorse, S., Roulleau, P., Figueiredo, S., & Benhamou, D. (2018). Impact of learners' role (active participant-observer or observer only) on learning outcomes during high-fidelity simulation	Frequency of closed-loop

	sessions in anaesthesia: A single center, prospective and randomised study. *Anaesthesia Critical Care & Pain Medicine*, *37*(5), 417–422. https://doi.org/10.1016/j.accpm.2017.11.016	communication not reported
6	Boet, S., Bould, M. D., Sharma, B., Revees, S., Naik, V. N., Triby, E., & Grantcharov, T. (2013). Within-Team Debriefing Versus Instructor-Led Debriefing for Simulation-Based Education: A Randomized Controlled Trial. *Annals of Surgery*, *258*(1), 53–58. https://doi.org/10.1097/SLA.0b013e31829659e4	Frequency of closed-loop communication not reported
7	Bragard, I., Farhat, N., Seghaye, M.-C., Karam, O., Neuschwander, A., Shayan, Y., & Schumacher, K. (2019). Effectiveness of a High-Fidelity Simulation-Based Training Program in Managing Cardiac Arrhythmias in Children: A Randomized Pilot Study. *Pediatric Emergency Care*, *35*(6), 412–418. https://doi.org/10.1097/PEC.0000000000000931	Frequency of closed-loop communication not reported
8	Bragard, I., Seghaye, M.-C., Farhat, N., Solowianiuk, M., Saliba, M., Etienne, A.-M., & Schumacher, K. (2018). Implementation of a 2-Day Simulation-Based Course to Prepare Medical Graduates on Their First Year of Residency. *Pediatric Emergency Care*, *34*(12), 857–861. https://doi.org/10.1097/PEC.0000000000000930	Not an RCT
9	Brindley, P. G., & Reynolds, S. F. (2011). Improving verbal communication in critical care medicine. *Journal of critical care*, *26*(2), 155–159. https://doi.org/10.1016/j.jcrc.2011.03.004	Not a primary study
10	Butchibabu, A., Sparano-Huiban, C., Sonenberg, L., & Shah, J. (2016). Implicit Coordination Strategies for Effective Team Communication. *Human Factors*, *58*(4), 595–610. https://doi.org/10.1177/0018720816639712	Not an RCT
11	Bruppacher, H. R., Alam, S. K., LeBlanc, V. R., Latter, D., Naik, V. N., Savoldelli, G. L., Mazer, C. D., Kurrek, M. M., & Joo, H. S. (2010). Simulation-based training improves physicians' performance in patient care in high-stakes clinical setting of cardiac surgery. *Anesthesiology*, *112*(4), 985–992. https://doi.org/10.1097/ALN.0b013e3181d3e31c	Frequency of closed-loop communication not reported
12	Capella, J., Smith, S., Philp, A., Putnam, T., Gilbert, C., Fry, W., Harvey, E., Wright, A., Henderson, K., Baker, D., Ranson, S., & Remine, S. (2010). Teamwork training improves the clinical care of trauma patients. *Journal of surgical education*, *67*(6), 439–443. https://doi.org/10.1016/j.jsurg.2010.06.006	Not an RCT
13	Carbo, A. R., Tess, A. V., Roy, C., & Weingart, S. N. (2011). Developing a high-performance team training framework for internal medicine residents: the ABC'S of teamwork. *Journal of patient safety*, *7*(2), 72–76. https://doi.org/10.1097/PTS.0b013e31820dbe02	Not an RCT
14	Cheng, A., Hunt, E. A., Donoghue, A., Nelson-McMillan, K., Nishisaki, A., Leflore, J., Eppich, W., Moyer, M., Brett-Fleegler, M., Kleinman, M., Anderson, J., Adler, M., Braga, M., Kost, S., Stryjewski, G., Min, S., Podraza, J., Lopreiato, J., Hamilton, M. F., Stone, K., … Nadkrni, V. M. (2013). Examining pediatric resuscitation education using simulation and scripted debriefing: a multicenter randomized trial. *JAMA pediatrics*, *167*(6), 528–536. https://doi.org/10.1001/jamapediatrics.2013.1389	Frequency of closed-loop communication not reported

15	Draycott, T., Sibanda, T., Owen, L., Akande, V., Winter, C., Reading, S., & Whitelaw, A. (2006). Does training in obstetric emergencies improve neonatal outcome? *BJOG : an International Journal of Obstetrics and Gynaecology*, *113*(2), 177–182. https://doi.org/10.1111/j.1471-0528.2006.00800.x	Not an RCT
16	Eppich, W. J., Adler, M. D., & McGaghie, W. C. (2006). Emergency and critical care pediatrics: use of medical simulation for training in acute pediatric emergencies. *Current opinion in pediatrics*, *18*(3), 266–271. https://doi.org/10.1097/01.mop.0000193309.22462.c9	Not a primary study
17	Figueroa, M. I., Sepanski, R., Goldberg, S. P., & Shah, S. (2013). Improving teamwork, confidence, and collaboration among members of a pediatric cardiovascular intensive care unit multidisciplinary team using simulation-based team training. *Pediatric cardiology*, *34*(3), 612–619. https://doi.org/10.1007/s00246-012-0506-2	Not an RCT
18	Flin, R., & Maran, N. (2004). Identifying and training non-technical skills for teams in acute medicine. *Quality & safety in health care*, *13 Suppl 1*(Suppl 1), i80–i84. https://doi.org/10.1136/qhc.13.suppl_1.i80	Not an RCT
19	Friedman, D., Zaveri, P., & O'Connell, K. (2010). Pediatric mock code curriculum: improving resident resuscitations. *Pediatric emergency care*, *26*(7), 490–494. https://doi.org/10.1097/PEC.0b013e3181e5bf34	Not an RCT
20	Fuhrmann, L., Perner, A., Klausen, T. W., Østergaard, D., & Lippert, A. (2009). The effect of multi-professional education on the recognition and outcome of patients at risk on general wards. *Resuscitation*, *80*(12), 1357–1360. https://doi.org/10.1016/j.resuscitation.2009.07.002	Not an RCT
21	Gilfoyle, E., Gottesman, R., & Razack, S. (2007). Development of a leadership skills workshop in paediatric advanced resuscitation. *Medical teacher*, *29*(9), e276–e283. https://doi.org/10.1080/01421590701663287	Not an RCT
22	Gillman, L. M., Brindley, P., Paton-Gay, J. D., Engels, P. T., Park, J., Vergis, A., & Widder, S. (2016). Simulated Trauma and Resuscitation Team Training course-evolution of a multidisciplinary trauma crisis resource management simulation course. *American journal of surgery*, *212*(1), 188–193.e3. https://doi.org/10.1016/j.amjsurg.2015.07.024	Not an RCT
23	Hicks, C. W., Rosen, M., Hobson, D. B., Ko, C., & Wick, E. C. (2014). Improving safety and quality of care with enhanced teamwork through operating room briefings. *JAMA surgery*, *149*(8), 863–868. https://doi.org/10.1001/jamasurg.2014.172	Not a primary study
24	Hunt, E. A., Walker, A. R., Shaffner, D. H., Miller, M. R., & Pronovost, P. J. (2008). Simulation of in-hospital pediatric medical emergencies and cardiopulmonary arrests: highlighting the importance of the first 5 minutes. *Pediatrics*, *121*(1), e34–e43. https://doi.org/10.1542/peds.2007-0029	Not an RCT
25	Hunt, E. A., Shilkofski, N. A., Stavroudis, T. A., & Nelson, K. L. (2007). Simulation: translation to improved team	Not a primary study

	performance. *Anesthesiology clinics*, *25*(2), 301–319. https://doi.org/10.1016/j.anclin.2007.03.004	
26	Hunziker, S., Bühlmann, C., Tschan, F., Balestra, G., Legeret, C., Schumacher, C., Semmer, N. K., Hunziker, P., & Marsch, S. (2010). Brief leadership instructions improve cardiopulmonary resuscitation in a high-fidelity simulation: a randomized controlled trial. *Critical care medicine*, *38*(4), 1086–1091. https://doi.org/10.1097/CCM.0b013e3181cf7383	Frequency of closed-loop communication not reported
27	Jankouskas, T., Bush, M. C., Murray, B., Rudy, S., Henry, J., Dyer, A. M., Liu, W., & Sinz, E. (2007). Crisis resource management: evaluating outcomes of a multidisciplinary team. *Simulation in healthcare : journal of the Society for Simulation in Healthcare*, *2*(2), 96–101. https://doi.org/10.1097/SIH.0b013e31805d8b0d	Not an RCT
28	Lingard, L., Whyte, S., Espin, S., Baker, G. R., Orser, B., & Doran, D. (2006). Towards safer interprofessional communication: constructing a model of "utility" from preoperative team briefings. *Journal of interprofessional care*, *20*(5), 471–483. https://doi.org/10.1080/13561820600921865	Qualitative
29	Miller, D., Crandall, C., Washington, C., 3rd, & McLaughlin, S. (2012). Improving teamwork and communication in trauma care through in situ simulations. *Academic emergency medicine : official journal of the Society for Academic Emergency Medicine*, *19*(5), 608–612. https://doi.org/10.1111/j.1553-2712.2012.01354.x	Not an RCT
30	Morey, J. C., Simon, R., Jay, G. D., Wears, R. L., Salisbury, M., Dukes, K. A., & Berns, S. D. (2002). Error reduction and performance improvement in the emergency department through formal teamwork training: evaluation results of the MedTeams project. *Health services research*, *37*(6), 1553–1581. https://doi.org/10.1111/1475-6773.01104	Not an RCT
31	Parsons, J. R., Crichlow, A., Ponnuru, S., Shewokis, P. A., Goswami, V., & Griswold, S. (2018). Filling the Gap: Simulation-based Crisis Resource Management Training for Emergency Medicine Residents. *The western journal of emergency medicine*, *19*(1), 205–210. https://doi.org/10.5811/westjem.2017.10.35284	Not an RCT
32	Risser, D. T., Rice, M. M., Salisbury, M. L., Simon, R., Jay, G. D., & Berns, S. D. (1999). The potential for improved teamwork to reduce medical errors in the emergency department. The MedTeams Research Consortium. *Annals of emergency medicine*, *34*(3), 373–383. https://doi.org/10.1016/s0196-0644(99)70134-4	Not an RCT
33	Saavedra, H. R., Turner, J. S., & Cooper, D. D. (2018). Use of Simulation to Improve the Comfort of Pediatric Residents Managing Critically Ill Emergency Department Patients. *Pediatric emergency care*, *34*(9), 633–635. https://doi.org/10.1097/PEC.0000000000001596	Not an RCT
34	Savoldelli, G. L., Naik, V. N., Park, J., Joo, H. S., Chow, R., & Hamstra, S. J. (2006). Value of debriefing during simulated crisis management: oral versus video-assisted oral	Frequency of closed-loop

	feedback. *Anesthesiology*, *105*(2), 279–285. https://doi.org/10.1097/00000542-200608000-00010	communication not reported
35	Shamim Khan, M., Ahmed, K., Gavazzi, A., Gohil, R., Thomas, L., Poulsen, J., Ahmed, M., Jaye, P., & Dasgupta, P. (2013). Development and implementation of centralized simulation training: evaluation of feasibility, acceptability and construct validity. *BJU international*, *111*(3), 518–523. https://doi.org/10.1111/j.1464-410X.2012.11204.x	Not an RCT
36	Sweeney, L. A., Warren, O., Gardner, L., Rojek, A., & Lindquist, D. G. (2014). A simulation-based training program improves emergency department staff communication. *American journal of medical quality : the official journal of the American College of Medical Quality*, *29*(2), 115–123. https://doi.org/10.1177/1062860613491308	Not an RCT
37	Thomas, S. M., Burch, W., Kuehnle, S. E., Flood, R. G., Scalzo, A. J., & Gerard, J. M. (2013). Simulation training for pediatric residents on central venous catheter placement: a pilot study. *Pediatric critical care medicine : a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies*, *14*(9), e416–e423. https://doi.org/10.1097/PCC.0b013e31829f5eda	Not an RCT
38	Tofil, N. M., Lee White, M., Manzella, B., McGill, D., & Zinkan, L. (2009). Initiation of a pediatric mock code program at a children's hospital. *Medical teacher*, *31*(6), e241–e247. https://doi.org/10.1080/01421590802637974	Not an RCT
39	Watters, C., Reedy, G., Ross, A., Morgan, N. J., Handslip, R., & Jaye, P. (2015). Does interprofessional simulation increase self-efficacy: a comparative study. *BMJ open*, *5*(1), e005472. https://doi.org/10.1136/bmjopen-2014-005472	Not an RCT
40	Wong, A. H., Gang, M., Szyld, D., & Mahoney, H. (2016). Making an "Attitude Adjustment": Using a Simulation-Enhanced Interprofessional Education Strategy to Improve Attitudes Toward Teamwork and Communication. *Simulation in healthcare : journal of the Society for Simulation in Healthcare*, *11*(2), 117–125.	Not an RCT

Total number of studies not available in full text (n=2)

#	Excluded studies	Reasons (found)
1	Complete the 'circle' when checking back. (2009). *ED Nursing, 12*(12), 139-.	Full-text not available
2	Curlis, J. Improving Closed Loop Communication in the Ipswich Emergency Department.	No publication details

8.4 Supplementary appendix 4: Characteristics of included studies and RoB

(Article)

General/reference information	
Source	
Country	

Study characteristics	**Description**
Study design	
Population	**Sample size:** **Occupational or educational roles:**
Intervention	**Approach:** **Duration:** **Simulated/real conditions**
Comparison	
Study results	**Outcome:** **Measurement methods:** **Data:**

(Lopez de Alda et al., 2021)

General/reference information	
Source	Lopez de Alda JX, Patel N, McNinch N, Ahmed RA. A Blindfolded Pediatric Trauma Simulation and Its Effect on Communication and Crisis Resource Management Skills. Cureus. 2021;13(11):e19484.
Country	USA

Study characteristics	**Description**
Study design	Randomized controlled trial
Population	**Total sample size:** N = 28 (blindfolded N = 15, non-blindfolded N = 13)
	Occupational or educational roles: Emergency medicine (EM) residents and pediatric EM fellows that had completed their Advanced Trauma Life support (ATLS) certification
Intervention	**Approach:** Participants in the intervention group led one pediatric trauma scenario simulation, blindfolded. Comprising the rest of the trauma team were three embedded participant nurses assigned to specific tasks. Different nurses participated on different simulation days. <u>Before the scheduled nursing staff participated, they:</u> • were provided the scenario to review • were instructed on appropriate CLC and how to communication with the leaders during the scenario <u>At the beginning of each simulation day, the scheduled nursing staff:</u> • partook in a rehearsal simulation to acclimate them to mannequins, the scenario, and expected communication

	Prior to starting the simulation, each participant from both the intervention and comparison group: • were shown a standardized PowerPoint presentation on CLC strategies • watched a video demonstrating a sample blindfolded simulation Following each simulation, participants from both the intervention and comparison group: • underwent a debriefing session
	Duration: Overall length (NI), simulation (10 minutes), debriefing session (10 minutes)
	Simulated/real conditions: Simulated conditions
Comparison	Participants in the comparison group led one pediatric trauma scenario simulation, non-blindfolded.
Study results (primary outcome)	Frequency of CLC **Outcome:** Frequency of CLC (presented as odds ratio and CI) **Measurement method:** Three independent faculty members reviewed recorded scenarios and counted the number of loop closures. For a loop to be considered closed, all CLC steps 1-4 had to be completed and agreed upon by at least two out of three reviewers. The likelihood of closing the loop during monitor placement in blindfolded groups compared to non-blindfolded groups was calculated. **Data:** OR = 13.7 (1.4–133.8).
Study results (secondary outcomes)	Medical errors (communication) **Outcome:** Missed CLC steps (presented as odds ratio and CI) **Measurement method:** Three independent faculty members reviewed recorded scenarios and counted the frequency of each step (1, 2, 3, and 4) of CLC. The number value agreed upon by at least two out of three reviewers was used. The likelihood of failing to complete step 3 and 4 during monitor placement in non-blinded groups compared to blindfolded groups was calculated. **Data:** OR step 3 = 6.2 (1.2-32.0), OR step 4 = 14.6 (2.2-97.6) Task performance **Outcome:** Completion of critical actions **Measurement method:** Three independent faculty members reviewed recorded scenarios for completion of critical actions. Not described which (or if) calculations were made for this outcome specifically. **Data:** No results or data presented. Task performance **Outcome:** Self-reported performance **Measurement method**: Participants were provided NASA Task Load Index (NASA-TLX) forms for completion post-participation. Not described which calculations were made for this outcome specifically. **Data**: Researchers found a non-significant difference, but supporting data is not presented. Task performance **Outcome:** CRM performance

		Measurement method: Emergency medicine, pediatric medicine, and simulation faculty members completed CRM forms after each scenario was completed. Although not explicitly stated, it is reasonable to assume that the Ottawa GRS scale was used to measure CRM performance in relation to the blindfolding effect, considering its use to measure CRM performance in relation to years of training. Not described which calculations were made for this outcome specifically. **Data**: Researchers found a non-significant difference, but supporting data is not presented.
Risk of bias assessment		
Entry	**Judgment**	**Support for judgement**
Random sequency generation (selection bias)	**Low risk**	*"Participants were randomized into groups…via a random number draw, with odd numbers representing blindfolded scenarios and even numbers representing non-blindfolded scenarios"* (p. 2). Comment: The random component in the sequence generation process is described.
Allocation concealment (selection bias)	**Unclear risk**	Comment: Allocation concealment not addressed. Insufficient information to permit judgement on low/high risk.
Blinding of participants and personnel (performance bias)	**High risk**	*"Prior to starting the simulation, each participant was shown a standardized PowerPoint presentation reviewing CLC strategies and then a video demonstrating a sample blindfolded simulation scenario." (p. 2)* Comment: Blinding not addressed for participants, scheduled nurses or involved instructors, but participants' introduction to blindfolding and CLC during pre-simulation preparations most certainly provided them with clues about the upcoming simulations' content and objectives, enabling them to easily assume their group assignment based on whether they were blindfolded or not, and intentionally or unintentionally perform accordingly.
Blinding of outcome assessment: self-reported performance (detection bias)	**High risk**	Comment: Due to this outcome being measured by the participants themselves, measurements have likely been influenced by the lack of blinding.
Blinding of outcome assessment: objective	**High risk**	Comment: Blinding of outcome assessment not addressed. Although no information clarifies whether the recordings were audio- or/and video-based, observations must have been

outcomes (detection bias)		required to assess CRM performance and completion of critical actions since these outcomes measure performance beyond mere verbal communication, enabling assessors to identify participants' group assignment from visual cues. As the methods used to assess the various objective outcomes appear to have involved some degree of outcome assessors' interpretation and judgement, there is a high risk that outcomes were influenced by the lack of blinding.
Incomplete outcome data: self-reported performance (attrition bias)	**Unclear**	Comment: The number of participants included in the statistical analysis for this particular outcome is not reported, and the authors do not mention loss to follow-up. Insufficient reporting of attrition/exclusions to permit judgement on low/high risk.
Incomplete outcome data: objective outcomes (attrition bias)	**Unclear**	Comment: The number of participants included in the statistical analysis for this particular outcome is not reported, and the authors do not mention loss to follow-up. Insufficient reporting of attrition/exclusions to permit judgement on low/high risk.
Selective reporting (reporting bias)	**High risk**	*"No significant differences were noted in CRM scores or NASA TLX scores"* (p. 4) Comment: No study protocol mentioned or found. The study fails to (1) clearly specify in the methods section which outcomes are considered primary and secondary, and (2) report results and outcome data for all outcomes assessed. The effects of blindfolding on CRM performance and self-reported performance are reported as non-significant, but without data. Results are not reported for the completion of critical actions.
Other biases	**Low risk**	Comment: The study appears to be free of other sources of bias.
Summary assessments of risk of bias (for data synthesis)		
Outcome	**Judgment**	**Support for judgement**
Self-reported performance	**High risk**	High risk of bias for one or more entries
Objective outcomes	**High risk**	High risk of bias for one or more entries

(Hughes et al., 2020)

General/reference information	
Source	Hughes KE, Hughes PG, Cahir T, Plitt J, Ng V, Bedrick E, et al. Advanced closed-loop communication training: the blindfolded resuscitation. BMJ Simulation & Technology Enhanced Learning. 2020;6(4):235-8.
Country	USA

Study characteristics	Description
Study design	Randomized controlled trial
Population	**Sample size:** N = 34 (blindfolded N = 17, non-blindfolded N = 17)
	Occupational or educational roles: Emergency medicine (EM) and EM/pediatric dual resident physicians
Intervention	**Approach:** Participants in the intervention group led two simulations; one adult resuscitation simulation and one pediatric resuscitation simulation, blindfolded. Comprising the rest of the resuscitation team were three embedded standardized participants for each participant. <u>Immediately before the simulations, participants from both the intervention and comparison group:</u> • watched a video demonstrating excellent and poor CLC • were instructed that orders would only be executed if communication was in perfect CLC format • were given an orientation to the manikin and simulation lab <u>After completing both simulations, participants from both the intervention and comparison group:</u> • partook in a one-on-one faculty debriefing
	Duration: Overall length (NI), adult simulation scenario (8 minutes), pediatric simulation scenario (8 minutes), debriefing (NI)
	Simulated/real conditions: Simulated conditions
Comparison	Participants in the comparison group led two simulations; one adult resuscitation simulation and one pediatric resuscitation simulation, non-blindfolded.
Study results (primary outcome)	<u>Frequency of CLC</u> **Outcome:** Frequency of perfect CLC use (presented as mean value and CI) **Measurement method:** Video-recorded simulations were reviewed by a single rater for frequency of CLC. The mean frequency was calculated for both the blindfolded group and the non-blindfolded group and tested for a statistically significant difference. **Data:** Blindfolded group: 31.7 (29.34-34.1), non-blindfolded group: 24.6 (21.5-27.7), p-value not presented.
Study results (secondary outcomes)	<u>Task performance</u> **Outcome:** Self-reported performance (presented as mean values and std. error)

	Measurement method: Participants used the NASA-TLX included in the post-surveys to assess their own perceived workload, including performance. Mean scores for self-reported performance were calculated for both the blindfolded group and non-blindfolded group and tested for a statistically significant difference. **Data:** Blindfolded group: 12.29 (1.55), non-blindfolded group: 12.06 (1.16), p = 0.78 <u>Task performance</u> **Outcome:** CRM performance **Measurement method:** One internal faculty member at the institution and two fellowship-trained EM physician raters reviewed video-recorded simulations and assessed CRM performance using the Ottawa GRS Scale. Mean Ottawa GRS scores were calculated and tested for a statistically significant difference. **Data:** Researchers found no significant difference, but supporting data is not presented.	
Risk of bias assessment		
Entry	**Judgment**	**Support for judgement**
Random sequency generation (selection bias)	**Unclear risk**	*"Participants were block randomized..."* (p. 235) <u>Comment:</u> Participants were randomly allocated, but the random component in the process of selecting the blocks is not specified. Insufficient information to permit judgement on low/high risk.
Allocation concealment (selection bias)	**Unclear risk**	<u>Comment:</u> Allocation concealment not addressed. Insufficient information to permit judgement on low/high risk.
Blinding of participants and personnel (performance bias)	**Unclear risk**	*"Participants watched internally made videos… demonstrating excellent and poor CLC. All participants were clearly instructed in the simulation prebriefed that orders would only be executed if communication was in perfect CLC format; non-verbal direction would be unsuccessful"* (p. 236) *"The resuscitation team was comprised of three embedded standardised participants for each participant"* (p. 236) <u>Comment:</u> Blinding not addressed for participants, standardized embedded participants, or the instructors involved. Participants were provided instructions and video-material on CLC prior to simulation, but no introduction to blindfolding or other information that would enable them to easily assume their group assignment. Unclear whether the

		instructors involved in the actual simulations were blinded. Also, if all participants shared the same embedded standardized participants (although this is unclear from the sentence above), the standardized participants would not have been blinded to allocated interventions. Insufficient information to permit judgement on low/high risk.
Blinding of outcome assessment: Self-reported performance (detection bias)	**Unclear risk**	Comment: Blinding of outcome assessment not addressed. Insufficient information to permit judgement on low/high risk.
Blinding of outcome assessment: objective outcomes (detection bias)	**High risk**	*"…two simulation fellowship-trained EM physician raters who were blind to participant identity and level of training and a third who was an internal faculty member at the institution."* (p. 236) Comment: No information is provided regarding blinding raters to allocated interventions, only to participants identity and level of training. Recordings were video-based, enabling assessors to identify participants' group assignment from visual cues (being blindfolded or not). As the methods used to assess the various objective outcomes appear to involve some degree of outcome assessors' interpretation and judgement, there is a high risk that outcomes were influenced by the lack of blinding.
Incomplete outcome data: perceived performance (attrition bias)	**Unclear risk**	Comment: The number of participants included in the statistical analysis for this particular outcome is not reported, and the authors do not mention loss to follow-up. Insufficient reporting of attrition/exclusions to permit judgement on low/high risk.
Incomplete outcome data: objective outcomes (detection bias)	**Unclear risk**	Comment: The number of participants included in the statistical analysis for this particular outcome is not reported, and the authors do not mention loss to follow-up. Insufficient reporting of attrition/exclusions to permit judgement on low/high risk.
Selective reporting (reporting bias)	**High risk**	*"When evaluating the mean Ottawa GRS scores… no category showed significant differences between main effects."* (p. 236) *"Frequency of closed loop communication significantly increased between the blindfolded group… and nonblindfolded group…»* (p. 236)

		Comment: No study protocol mentioned or found. The effect of blindfolding on CRM performance is reported as non-significant, but without data. The difference in frequency of CLC is reported as significant, although no p-value statistically confirms this statement.
Other biases	**Low risk**	Comment: The study appears to be free of other sources of bias.
Summary assessments of risk of bias (for data synthesis)		
Outcome	**Judgment**	**Support for judgement**
Self-reported performance	**High risk**	Comment: High risk of bias for one or more entries.
Objective outcomes	**High risk**	Comment: High risk of bias for one or more entries

(Scicchitano et al., 2021)

General/reference information	
Source	Scicchitano E, Stark P, Koetter P, Michalak N, Zurca AD. Blindfolding Improves Communication in Inexperienced Residents Undergoing ACLS Training. Journal of graduate medical education. 2021;13(1):123-7.
Country	USA

Study characteristics	**Description**
Study design	Randomized controlled trial
Population	**Sample size:** N = 87 (blindfolded N = 46, non-blindfolded N = 41)
	Occupational or educational roles: Medical interns undergoing advanced cardiovascular life support (ACLS) training at a single academic medical center
Intervention	**Approach:** Participants in the intervention groups functioned as team leaders in two simulations (rotating into the leadership role in turns); once during a practice session, blindfolded, and once during a testing session, non-blindfolded. Also referred to as blindfolded training (BT). Prior to the practice sessions, participants from both the intervention and comparison groups: • watched a presentation about resuscitation team dynamics provided by the American Heart Association (AHA) At the beginning of the practice sessions, participants from both the intervention and comparison groups: • were given standardized instructions. The blindfolded groups were instructed on how they would be deprived of sight During the practice sessions only, participants from both the intervention and comparison groups:

	• were allowed to view the ACLS algorithms
	Duration: NI
	Simulated/real conditions: Simulated conditions
Comparison	Participants in the comparison groups functioned as a team leader in two simulations (rotating into the leadership role in turns); once during practice session and once during a testing session, both non-blindfolded. Also referred to as standard training (ST).
Study results (primary outcome)	Frequency of CLC: **Outcome:** Number of complete closed-loop communication (presented as mean value and CI) **Measurement method:** Verbal communications were coded into five different types of communication events, including the use of complete closed-loop communication. Three study members coded the first two videos, two study members double coded the next 20% of recordings, and one of two study members coded the remaining videos. Flagged communications were discussed between the three study members to achieve consensus. The mean number of complete closed-loop communications across BT groups and across ST groups was calculated and tested for a statistically significant difference. **Data:** BT groups = 20.3 (18.8-21.8), ST groups = 16.6 (14.8-18.4), p = 0.002
Study results (secondary outcomes)	Task performance **Outcome:** Time from cardiac arrest to initiation of chest compression (presented as mean value and CI) **Measurement method:** Time measured in seconds. The mean number of seconds from cardiac arrest to chest compressions across BT groups and across ST groups was calculated and tested for a statistically significant difference. **Data:** BT groups = 13.96 (10.2-17.8), ST groups = 13.8 (8.7-19), p = 0.96 Task performance **Outcome:** Time from cardiac arrest to defibrillation (presented as mean value and CI) **Measurement method:** Time measured in seconds. The mean number of seconds from cardiac arrest to defibrillation across BT groups and across ST groups was calculated and tested for a statistically significant difference. **Data:** BT groups = 55.9 (44.8-67.1), ST groups = 50.6 (38.9-62.3), p = 0.51 Task performance **Outcome:** Time from cardiac arrest to first dose of epinephrine (presented as mean value and CI) **Measurement method:** Time measured in seconds. The mean number of seconds from cardiac arrest to first dose of epinephrine across BT groups and across ST groups was calculated and tested for a statistically significant difference. **Data:** BT groups = 105.8 (86.4-125.3), ST groups = 108.9 (89-128.9), p = 0.82

Notes	*For all outcomes above, video- and audio-recordings of the individual testing scenarios were viewed and evaluated. For verbal communication, 3 study team members were the outcome assessors. It remains unclear whether the same assessors also measured performance.*	
Risk of bias assessment		
Entry	**Judgment**	**Support for judgement**
Random sequency generation (selection bias)	**Unclear risk**	*"The interns were divided into groups, and groups were randomized..."* (p. 124) Comment: Pre-formed groups of participants were randomized, but the random component in the sequence generation process is not described. Insufficient information to permit judgement on low/high risk.
Allocation concealment (selection bias)	**Unclear risk**	Comment: Allocation concealment not addressed. Insufficient information to permit judgement on low/high risk.
Blinding of participants and personnel (performance bias)	**High risk**	*"… instructors received written and verbal orientations to the study and study goals and taught practice sessions for both groups."* (p. 124) *«Instructors for the testing sessions were blinded to whether groups had been randomized to BT or ST»* (p. 124) *«Participants were blinded to the objectives and outcomes of the study»* (p. 124) Comment: Unclear whether participants' blinding to the study's objectives and outcomes means that they were also unable to identify their group assignment. Even though the blinding of instructors for the testing sessions, from which the outcomes were measured, was probably considered most important, the instructors for the practice sessions appear to not have been blinded. This may have influenced their instructional and teaching delivery, leading to different experiences among participant groups prior to the upcoming testing sessions, potentially affecting participants' performance during the actual testing phase.
Blinding of outcome assessment: Self-reported performance (detection bias)	**N/A**	Comment: The study did not address this outcome.
Blinding of outcome	**Low risk**	

assessment: objective outcomes (detection bias)		*«… coded by 3 study team members blinded to which type of training the interns had randomized… to ensure concordant definitions for the different types of communication. »* (p. 124) Comment: Blinding of outcome assessment was ensured for verbal communication. Although it remains unclear whether the same blinded assessor(s) or someone else documented performance, it is unlikely that blinding could have been broken as no participants were blindfolded on the video-recordings used for evaluation. Besides, the direct measure of performance in terms of time-to-task completion (in seconds), makes this outcome less prone to subjective interpretation.
Incomplete outcome data: self-reported performance (attrition bias)	**N/A**	Comment: The study did not address this outcome.
Incomplete outcome data: objective outcomes (attrition bias)	**Low risk**	*"Two interns were excluded from analysis due to issues with video and data acquisition, both of which were in the ST group. In total, 85 of 87 (98%) interns were included in the analysis, of which 39 underwent ST and 46 underwent BT." (p. 125)* Comment: Despite an uneven proportion of incomplete outcome data across groups, data are missing for only 2% of all participants, and the reason appears to be related to technical difficulties rather than their true outcomes.
Selective reporting (reporting bias)	**Low risk**	Comment: No study protocol mentioned or found, but results and outcome data for all outcomes listed in the method section of the article have been reported.
Other biases	**Low risk**	Comment: The study appears to be free of other sources of bias.
Summary assessments of risk of bias (for data synthesis)		
Outcome	**Judgment**	**Support for judgement**
Self-reported performance	**N/A**	Comment: The study did not address this outcome.
Objective outcomes	**High risk**	Comment: High risk of bias for one or more entries

(Buyck et al., 2019)

General/reference information	
Source	Buyck M, Manzano S, Haddad K, Moncousin AC, Galetto-Lacour A, Blondon K, et al. Effects of blindfold on leadership in pediatric resuscitation simulation: a randomized trial. Front Pediatr. 2019;7:10. doi:10. 3389/fped.2019.00010.
Country	Switzerland

Study characteristics	**Description**
Study design	Randomized controlled trial
Population	**Sample size:** N = 48 (blindfolded N = 24, non-blindfolded N = 24)
	Occupational or educational roles: Pediatric emergency fellows, pediatric emergency residents, and pediatric emergency registered nurses (recruited among the staff of the pediatric emergency department)
Intervention	**Approach:** Participants in the intervention group took part in five high-fidelity simulation-based pediatric resuscitation scenarios (A, B, C, D and E). The team leader of the group was blindfolded during simulations B, C, and D, but not A and E. This group was called the blindfold group (BG). Participants from both the intervention and comparison group: • underwent a standardized debrief focusing on the non-technical skills of the leader after simulations B, C, and D • took part in simulations A and E as a pre- and post-test, without any debrief • underwent identical scenarios for simulations B, C and D, while scenarios for simulations A and E were considered comparable between the intervention and comparison group.
	Duration: Overall length (4 hours), simulation (NI) debriefing session (maximum of 20 minutes x 3)
	Simulated/real conditions: Simulated conditions
Comparison	Participants in the comparison group took part in five different high-fidelity simulation-based pediatric resusciation scenarios, all with the team leader non-blindfolded. Referred to as the control group (CG).
Study results (primary outcome)	Frequency of CLC: **Outcome:** Number of complete communication loops (presented as median value and IQR) **Measurement method:** A communication loop was considered complete when three pre-specified verbal elements were present. The median value and IQR were calculated for both the blindfold group and the control group and tested for a statistically significant difference. **Data:** Blindfold group: 0 (-7;5), Control group: +3 (-1;4), p = 0.63
Study results (secondary outcomes)	Medical errors (communication) **Outcome:** Number of incomplete communication loops (presented as median value and IQR) **Measurement method:** The median value and IQR for the change in number of incomplete communication loops were calculated for both the

	blindfold group and the control group and tested for a statistically significant difference. **Data:** Blindfold group: -2 (-4;-1), Control group: 0 (-2;0), p = 0.05 <u>Task performance</u> **Outcome:** Time to cardio-pulmonary resuscitation (CPR) from cardiac arrest (presented as median value and IQR) **Measurement method:** Time measured in seconds. The median value and IQR for change in time from circulatory arrest to CPR were calculated for both the blindfold group and the control group and tested for a statistically significant difference. initiation. **Data:** Blindfold group: 190 (58;267), Control group: 61 (17;151), p = 0.15 <u>Task performance</u> **Outcome:** Number of pertinent reassessments (presented as median value and IQR) **Measurement method:** The median value and IQR for change the number of reassessments were calculated for both the blindfold group and the control group and tested for a statistically significant difference. **Data:** Blindfold group: 1 (-1;3), Control group 0 (-2;3), p = 0.57 <u>Task performance</u> **Outcome:** Progression of the Resuscitation team leader evaluation score (presented as percentage and IQR) **Measurement method:** Using the scale by Grant et al., all non-clinical skills except number 1 and 12 were scored from 0-30 points. The median value and IQR for the proportion of improvement Calgary score were calculated for both the blindfold group and the control group and tested for a statistically significant difference. **Data:** Blindfold group: 11.4% (8.0;18.9), Control group: 5.4% (0.0;8.6), p = 0.04	
Notes	*For all outcomes above, three simulation-based training experts who did not take part in the simulation assessed and scored the video recordings from simulations A and E.*	
Risk of bias assessment		
Entry	**Judgment**	**Support for judgement**
Random sequency generation (selection bias)	**Unclear risk**	*"They were assigned to 12 resuscitation teams…"* (p. 2) *"These 12 teams were randomly allocated following simple randomization procedure…"* (p. 2) <u>Comment:</u> Pre-formed groups of participants were randomized, but the random component in the sequence generation process have not been described. Insufficient information to permit judgement on low/high.

Allocation concealment (selection bias)	**Unclear risk**	*"… (sealed envelopes with a 50% chance) …" (p. 2)* Comment: Adequate concealment of allocations would require the envelopes to be opaque and sequentially numbered, in addition to being sealed, but these safeguards are not mentioned. Insufficient information to permit judgement on low/high.
Blinding of participants and personnel (performance bias)	**Unclear risk**	*"… the simulation instructor was not blinded to the allocation as he had to observe the simulations. To address this potential bias, our debriefings were strictly scripted and monitored by the primary investigator who was present for all simulations and debriefings."* (p. 5) Comment: The simulation instructor who gave debriefings was not blinded, but measures were taken to ensure that all the debriefings were conducted in the same way regardless of group allocation. Blinding not addressed for participants or instructors involved in the actual simulations. Insufficient information to permit judgement on low/high risk.
Blinding of outcome assessment: self-reported performance (detection bias)	**N/A**	Comment: The study did not address this outcome.
Blinding of outcome assessment: objective outcomes (detection bias)	**Low risk**	*«Simulations A and E video recordings were scored by three simulation-based training experts who did not take part to the simulations. The experts were blinded to the team allocations to CG or BG» (p. 3)* Comment: Blinding of outcome assessment was ensured, and unlikely to have been broken as experts did not partake in simulations, nor could have distinguished between groups based on visual cues due to no participants being blindfolded on the video-recordings used for evaluation.
Incomplete outcome data: self-reported performance (attrition bias)	**N/A**	Comment: The study did not address this outcome.
Incomplete outcome data: objective	**Unclear**	Comment: The number of participants included in the statistical analysis for this particular outcome is not reported, and the authors do not mention loss to follow-up. Insufficient

outcomes (detection bias)		reporting of attrition/exclusions to permit judgement on low/high risk.
Selective reporting (reporting bias)	**Low risk**	Comment: No study protocol mentioned or found, but results and outcome data for all outcomes described in the method section of the article have been reported.
Other biases	**Low risk**	Comment: The study appears to be free of other sources of bias.
Summary assessments of risk of bias (for data synthesis)		
Outcome	**Judgment**	**Support for judgement**
Self-reported performance	**N/A**	Comment: The study did not address this outcome.
Objective outcomes	**Unclear risk**	Comment: Unclear risk of bias for one or more entries

8.5 Supplementary appendix 5: GRADE assessment

Grade domain	Judgement	Rating
Methodological limitations	No more than a total of 197 participants were included across the four studies. The two studies largest in	Serious concerns

	participant size (n = 87 and n = 48) had low risk for blinding of outcome assessment, selective reporting, and other biases (Buyck et al., 2019; Scicchitano et al, 2021). Scicchitano et al. (2021) also had low risk of incomplete outcome data. Scicchitano et al. (2021) had no entries at high risk of bias, while Buyck et al. (2019) had high risk bias only for blinding of participants and personnel. The remaining studies (n = 28 and n = 34) had high risk of bias for either two or three of the following entries: blinding of participants and personnel, blinding of outcome assessment and selective reporting (Lopez de Alda et al., 2021; Hughes et al., 2020). Multiple entries in all four studies were rated as unclear due to insufficient information.	
Indirectness	All four studies provided research evidence directly linked to the question at hand. They all included participants either specialized in or focusing on emergency medical care, a blindfolding intervention, and a non-blindfolding as a comparative group.	No serious concerns
Imprecision	The total sample size including all studies were below 400, and therefore not sufficient to dismiss concerns about imprecision. The precision of results varied from reporting considerably wide 95%CIs to small p-values.	Serious concerns
Inconsistency	Three of the four studies show consistency in the direction of effect, favoring the blindfolding intervention, though the magnitude of effect sizes varied substantially from moderate to very large. The remaining study reports a contrasting no effect of blindfolding on the intervention group, and a statistically non-significant between-group difference. Also, important variations between studies were seen in terms of occupational roles, intervention components, measures and reporting of outcomes. Lopez de Alda et al. (2021) and Buyck et al. (2019) varied in terms of what they considered to be a complete closed-loop communication, and the other two studies provided no such criteria.	Serious concerns
Publication bias	The searches for identifying relevant studies were comprehensive.	No serious concerns