

BMJ Open How effective is teamwork really? The relationship between teamwork and performance in healthcare teams: a systematic review and meta-analysis

Jan B. Schmutz,¹ Laurenz L. Meier,² Tanja Manser³

To cite: Schmutz JB, Meier LL, Manser T. How effective is teamwork really? The relationship between teamwork and performance in healthcare teams: a systematic review and meta-analysis. *BMJ Open* 2019;9:e028280. doi:10.1136/bmjopen-2018-028280

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-028280>).

Received 29 November 2018
Revised 14 August 2019
Accepted 16 August 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of Communication Studies, Northwestern University, Evanston, Illinois, USA

²Department of Work and Organizational Psychology, University of Neuchâtel, Neuchâtel, Switzerland

³FHNW School of Applied Psychology, University of Applied Sciences and Arts Northwestern Switzerland, Olten, Switzerland

Correspondence to

Dr Jan B. Schmutz;
jan.schmutz@northwestern.edu

ABSTRACT

Objectives To investigate the relationship between teamwork and clinical performance and potential moderating variables of this relationship.

Design Systematic review and meta-analysis.

Data source PubMed was searched in June 2018 without a limit on the date of publication. Additional literature was selected through a manual backward search of relevant reviews, manual backward and forward search of studies included in the meta-analysis and contacting of selected authors via email.

Eligibility criteria Studies were included if they reported a relationship between a teamwork process (eg, coordination, non-technical skills) and a performance measure (eg, checklist based expert rating, errors) in an acute care setting.

Data extraction and synthesis Moderator variables (ie, professional composition, team familiarity, average team size, task type, patient realism and type of performance measure) were coded and random-effect models were estimated. Two investigators independently extracted information on study characteristics in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.

Results The review identified 2002 articles of which 31 were included in the meta-analysis comprising 1390 teams. The sample-sized weighted mean correlation was $r=0.28$ (corresponding to an OR of 2.8), indicating that teamwork is positively related to performance. The test of moderators was not significant, suggesting that the examined factors did not influence the average effect of teamwork on performance.

Conclusion Teamwork has a medium-sized effect on performance. The analysis of moderators illustrated that teamwork relates to performance regardless of characteristics of the team or task. Therefore, healthcare organisations should recognise the value of teamwork and emphasise approaches that maintain and improve teamwork for the benefit of their patients.

INTRODUCTION

May it be an emergency team in the trauma room, paramedics treating patients after an accident or a surgical team in the operating room, teams are ubiquitous in healthcare and must work across professional, disciplinary

Strengths and limitations of this study

- This systematic review evaluates available studies investigating the effectiveness of teamwork processes.
- Thirty-one studies have been included resulting in a substantial sample size of 1390 teams.
- The sample size of the primary studies included is usually low.
- For some subgroup analysis, the number of studies included was small.

and sectorial boundaries. Although the clinical expertise of individual team members is important to ensure high performance, teams must be capable of applying and combining the unique expertise of team members to maintain safety and optimal performance. In order for a team to be effective individual team members need to collaborate and engage in teamwork. Today, experts agree that effective teamwork anchors safe and effective care at various levels of the healthcare systems¹⁻⁴ leading to a relatively recent shift towards team research and training.⁵⁻⁷

Healthcare is an evidence-based field and therefore administrators and providers are seeking evidence in the literature concerning the impact of teamwork on performance outcomes like patient mortality, morbidity, infection rates or adherence to clinical treatment guidelines. Having a closer look at the literature investigating healthcare teams we find mixed and sometimes even contradicting results about the relationship between teamwork and clinical performance.⁸ Some studies find a large effect of teamwork on performance outcomes (eg, Carlson *et al*)⁹ while others report small or no relationships.^{10 11} This inconsistency arises due to several reasons. First, the conceptual and empirical literature examining teamwork is fragmented and research examining

teamwork effectiveness is spread across disciplines including medicine, psychology and organisation science. Therefore, researchers and practitioners often lack a common conceptual foundation for investigating teams and teamwork in healthcare. Second, research studies on teamwork in healthcare usually exhibit small sample sizes because of the challenges of recruiting actual professional teams and carefully balancing research with patient care priorities. Small sample sizes, however, increase the likelihood of reporting results that fail to represent true effect. Third, studies investigating healthcare teams often ignore important context variables of teams (eg, team composition and size, task characteristics, team environment) that likely influence the effect that teamwork has on clinical performance.^{12 13}

These inconsistencies in the teamwork literature may lead to confusion about the importance of teamwork in healthcare, thus giving voice to critics who hinder efforts to improve teamwork. We aim to address these problems with a meta-analytical study investigating the performance implications of teamwork. A meta-analytical approach moves beyond existing reviews on teamwork in healthcare^{8 14–17} and quantitatively tests if the widely advocated positive effect of teamwork on performance holds true. In addition, this approach allows us to investigate context variables as moderators that may influence the effect of teamwork on performance, meaning that this effect might be stronger or weaker under certain conditions. Previous meta-analyses^{18 19} focused mainly on the effectiveness of team trainings but not on the effect of teamwork itself. This meta-analysis will generate quantitative evidence to inform the relevance of future interventions, regulations and policies targeting teamwork in healthcare organisations.

In the following we will first establish an operational definition of teamwork, elaborate on relevant contextual factors and present our respective meta-analytical results and their interpretation.

Teams, teamwork and team performance

In order to clearly understand the impact of teamwork on performance it is necessary to provide a brief introduction to teams, teamwork and team performance. We define teams as identifiable social work units consisting of two or more people with several unique characteristics. These characteristics include (a) dynamic social interaction with meaningful interdependencies, (b) shared and valued goals, (c) a discrete lifespan, (d) distributed expertise and (e) clearly assigned roles and responsibilities.^{20 21} Based on this definition it becomes clear that teams must dynamically share information and resources among members and coordinate their activities in order to fulfil a certain task — in other words teams need to engage in *teamwork*.

Teamwork as a term is widely used and often difficult to grasp. However, we absolutely require a clear definition of teamwork especially for team trainings that target specific behaviours. *Teamwork* is a process that describes

interactions among team members who combine collective resources to resolve task demands (eg, giving clear orders).^{22 23} Teamwork or team processes can be differentiated from *taskwork*. *Taskwork* denotes a team's individual interaction with tasks, tools, machines and systems.²³ *Taskwork* is independent of other team members and is often described as *what* a team is doing whereas *teamwork* is *how* the members of a team are doing something with each other.²⁴ Therefore, *team performance* represents the accumulation of teamwork and taskwork (ie, what the team actually does).²⁵

Team performance is often described in terms of inputs, processes and outputs (IPO).^{22 26–28} *Outputs* like quality of care, errors or performance are influenced by team related *processes* (ie, teamwork) like communication, coordination or decision-making. Furthermore, these processes are influenced by various *inputs* like team members' experience, task complexity, time pressure and more. The IPO framework emphasises the critical role of team processes as the mechanism by which team members combine their resources and abilities, shaped by the context, to resolve team task demands. It has been the basis of other more advanced models^{27–29} but has also been criticised because of its simplicity.³⁰ However, it is still the most popular framework to date and helps to systematise the mechanisms that predict team performance and represents the basis for the selection of the studies included in our meta-analysis.

Contextual factors of teamwork effectiveness

Based on a large body of team research from various domains, we hypothesise that several contextual and methodological factors might moderate the effectiveness of teamwork, indicating that teamwork is more important under certain conditions.^{31 32} Therefore, we investigate several factors: (a) team characteristics (ie, professional composition, team familiarity, team size), (b) task type (ie, routine vs non-routine tasks), (c) two methodological factors related to patient realism (ie, simulated vs real) and (d) the type of performance measures used (ie, process vs outcome performance). In the following we discuss these potentially moderating factors and the proposed effects on teamwork.

Professional composition

We distinguished between interprofessional and uniprofessional teams. Interprofessional teams consist of members from various professions that must work together in a coordinated fashion.³³ Diverse educational paths in interprofessional teams may shape respective values, beliefs, attitudes and behaviours.³⁴ As a result team members with different backgrounds might perceive and interpret the environment differently and have a different understanding of how to work together. Therefore, we assume that explicit teamwork is especially important in interprofessional teams compared with uniprofessional teams.

Team familiarity

If team members have worked together, they are familiar with their individual working styles; and roles and responsibilities are usually clear. If a team works together for the first time, this potential lack of familiarity and clarity might make teamwork even more important. Therefore, we differentiate between *real teams* that also work together in their everyday clinical practice and *experiential teams* that only came together for study purposes.

Team size

Another factor that may moderate the relationship between teamwork and performance is team size. Since larger teams exhibit more linkages among members than smaller teams, they also face greater coordination challenges. Also, with increasing size teams have greater difficulty developing and maintaining role structures and responsibilities. For these reasons, we expect the influence of teamwork on clinical performance to be stronger in larger teams as compared with smaller teams.

Task type

Routine situations are characterised by repetitive and unvarying actions (eg, standard anaesthesia induction).³⁵ In contrast, non-routine situations exhibit more variation and uncertainty, requiring teams to be flexible and adaptive. Whereas team members mostly rely on pre-learned sequences during routine situations, during non-routine situations we assume that teamwork is more important in order for team members to resolve task demands.

Patient realism

Authors highlight the importance of using medical simulators in education.³⁶ Therefore, we investigate the realism used in a study (simulated vs real patients) as a potential methodological factor that influences the relationship between teamwork and performance. Studies conducted with medical simulators might be more standardised and less influenced by confounding variables than studies conducted with real patients. Therefore, results from simulation studies might show stronger relationships between the two variables. Further, using a simulator could cause individuals and teams to act differently than in real settings, thereby distorting the results. However, in the last decade high-fidelity simulators have become increasingly realistic, suggesting that the results from simulation studies generalise to real environments. Including realism as a contextual factor in our analysis will reveal if the effects of teamwork observed in simulation compare with real life settings. Better understanding would provide important insights about simulation use in teamwork studies.

Performance measures

As a second methodological factor, we expect that the type of performance measure used in a study influences the reported teamwork effectiveness. The literature usually differentiates between process-related and outcome-related aspects of performance.^{37 38} Process performance

measures are action-related aspects and refer to adequate behaviour during procedures (eg, adhering to guidelines), making them easier to assess. Outcome performance measures (eg, infection rates after operations) follow team actions, with assessment occurring later than process measures. Outcome performance measures suffer from several factors: greater sensitivity to confounding variables (eg, comorbidities), assessment challenges and greater difficulty linking team processes to outcomes. Looking at the predictors of the survival of cardiac arrest patients illustrates the difference between the two types of performance measures. The main predictors for the survival (ie, performance outcome) of a cardiac arrest patient are 'duration of the arrest' and 'age of the patient less than 70'.³⁹ Although a team delivers perfect basic life support (ie, high process performance) the patient can still die (ie, low outcome performance). Due to these methodological considerations, we expect that studies assessing process performance report a stronger relationship between teamwork and performance than studies assessing outcome performance.

METHODS

The study was conducted based on the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement⁴⁰ as well as established guidelines in social sciences.^{41 42} Through the combination of studies in the meta-analytical process, we will increase the statistical power and provide an accurate estimation of the true impact that teamwork has on performance.

Search strategy

We applied the following search strategy to select relevant papers: (a) an electronic search of the database PubMed (no limit was placed on the date of publication, last search 19 June 2018) using the keywords *teamwork*, *coordination*, *decision-making*, *leadership* and *communication* in combination with *patient safety*, *clinical performance*, the final syntax for PubMed is available (online supplementary file), (b) a manual backwards search for all references cited by eight systematic literature reviews that focus on teamwork or non-technical skills in various healthcare domains,^{8 15 17 43–47} (c) a manual backwards search for all references cited in studies we included in our meta-analysis, (d) a manual forward search using Web of Science to identify studies that cite the studies we included in our meta-analysis, (e) identification of relevant unpublished manuscripts via email from authors currently investigating medical teams using specific mailing lists.

Inclusion criteria

Studies were included if a construct complied to the definition of teamwork processes outlined in the introduction (eg, coordination, communication). In addition, studies needed to investigate the relationship between at least one teamwork process and a performance measure (eg, patient outcome). When studies reported multiple

estimates of the same relationship from the same sample (eg, between coordination and more than one indicator of performance), those correlations were examined separately only as appropriate for sub-analyses, but an average correlation was computed for all global meta-analyses of those relationships to maintain independence.⁴¹ We excluded articles investigating long-term care since the coordination of care for chronically ill patients has to consider the unique team task interdependencies in this setting.⁴⁸ Also, teams working together over longer periods of time are more likely to develop emergent states (eg, team cohesion) that influence how a specific team works together.²⁴ All articles included in this meta-analysis are listed in [tables 1 and 2](#).

For the criterion level of analysis, we included only effect sizes at the team level and not on an individual level. Therefore, the performance measure had to be clearly linked to a team. This approach aligns with research that strongly recommends against mixing levels of analysis in meta-analytical integrations.^{49 50}

Two reviewers independently screened titles and abstracts from articles yielded in the search. Afterwards full texts of all relevant articles were obtained and screened by the same two reviewers. Agreement was above 90%. Any disagreement in the selection process was resolved through consensus discussion.

Data extraction

With the help of a jointly developed coding scheme, studies were independently coded by one of the authors (JS) and another rater, both with a background in industrial psychology and human factors. Twenty per cent of the studies were rated by both coders. Inter-coder agreement was above 90%. Any disagreement was resolved through discussion. The data extracted comprised details of the authors and publication as well as important study characteristics and statistical relationships between a teamwork variable and performance ([table 2](#)).

Coding of team characteristics

The *professional composition* of teams was coded either as 'Interprofessional' if a team consisted of members from different professions (eg, nurses and physicians) or as 'Uniprofessional' if the members of the teams were of the same profession. *Team size* was coded as the number of members (average number if team size varied) of the investigated teams. Team familiarity was coded either as 'experimental' or 'real'. 'Real' indicates that the team members also worked together in their everyday clinical practice. 'Experimental' means that the teams only worked together during the study.

Coding of task characteristics

Task type was coded either as 'Routine task' or 'Non-routine task'. We defined 'Non-routine tasks' as unexpected events that require flexible behaviour often under time-pressure (eg, emergency situations). 'Routine tasks'

describe previously planned standard procedures (eg, standard anaesthesia induction, planned surgery).

Coding of methodological factors

Patient realism was either coded as 'Real patient' or 'Simulated patient'. 'Simulated patient' included a patient simulator (manikin) whereas 'Real patient' included real patients in clinical settings.

Clinical performance measures were coded either as 'Outcome performance' or 'Process performance'.^{38 51} 'Outcome performance' includes an outcome that is measured after the treatment process (eg, infection rate, mortality). We focused only on patient-related outcomes and not on team outcomes (eg, team satisfaction). 'Process performance' describes the evaluation of the treatment process and describes how well the process was executed (eg, adherence to guidelines through expert rating). Process performance measures are often based on official guidelines and extensive expert knowledge.⁵² Thus, we assumed that process performance closely relates to patient outcomes.

Statistical analysis

Different types of effect sizes (eg, OR, *F* values and *r*) have been reported in the original studies. We therefore converted the different effect sizes to a common metric, namely *r* using the formulas provided by Borenstein *et al*.⁵³ and Walker.⁵⁴ Moreover, some samples contained effect sizes of teamwork with two or more measures of performance. Because independence of the included effects sizes is required for a meta-analysis,^{41 55} we used Fisher's *z* score to average the multiple correlations from the same sample (scholars have suggested to convert *r* to Fisher's *z* scores, to average the *z*'s and then to back-transform it to *r*.⁵⁶ Using simple arithmetic average (ie, correlations will be summed and divided by the number of coefficients) is problematic because the distribution of *r* becomes negatively skewed as the correlation is larger than zero. As a result, the average *r* tends to underestimate the population correlation). The correlations were weighted for sample size. However, in contrast to many meta-analyses in social sciences, the correlations were not adjusted for measurement reliability. This is because information about the measurement reliability could not be compared (Kappa vs Cronbach's Alpha) or were not available at all for the majority of studies. Therefore, we report uncorrected, sample-size weighted mean correlation, its 95% CI, and the 80% credibility interval (CR). The CI reflects the accuracy of a point estimate and can be used to examine the significance of the effect size estimates, whereas the CR refers to the deviation of these estimates and informs us about the existence of possible moderators.

Random-effects models were estimated based on two considerations.⁵⁷ First, we expected study heterogeneity to be high given the different study design characteristics such as *patient realism* ('Real patient' vs 'Simulated patient'), *task type* ('Routine task' vs 'Non-routine task')

Table 1 Descriptions of study objectives, settings and description of teamwork process and outcome measures

Authors	Year	Main study objectives	Participants and setting	Teamwork process measure	Outcome measure
Amacher <i>et al</i> ⁶⁰	2017	To compare female and male rescuers in regard to cardiopulmonary resuscitation and leadership performance	Video observation of medical students managing cardiopulmonary resuscitation in a high-fidelity patient simulator	Structured observation of secure leadership statements within teams	Time until start chest compression hands-on time within first 180 s
Brogaard <i>et al</i> ⁶⁰	2018	To investigate the relationship between non-technical skills and clinical performance in obstetrical teams	Video observation of obstetrical teams (obstetricians, obstetrical nurse, anaesthesiologists) managing real-life emergencies (postpartum haemorrhage)	Assessment of non-technical skills using a behaviourally anchored rating scale (ATOP); Assessment of Obstetric Team Performance)	Checklist tool for clinical performance (TeamOBS-PPH)
Burtscher <i>et al</i> ⁸¹	2011	To investigate how team mental models and team monitoring behaviour interact to predict team performance in anaesthesia	Video observation of anaesthesia teams (residents, nurses) conducting a standard anaesthesia induction using a high-fidelity patient simulator	Structured observation of team monitoring behaviour	Checklist based expert rating
Burtscher <i>et al</i> ⁸²	2011	To investigate the relationship between adaptation of team coordination and clinical performance in response to a critical event	Video observation of anaesthesia teams (resident, nurse) conducting a standard anaesthesia induction including a critical event using a high-fidelity patient simulator	Structured observation of team coordination	Reaction time related to the critical event
Burtscher <i>et al</i> ⁸³	2010	To examine the role of anaesthesia teams' adaptive coordination in managing changing situational demands	Video observation of anaesthesia teams (residents, nurses, students) conducting standard anaesthesia inductions with non-routine events	Structured observation of team coordination	Checklist based expert rating
Carlson ⁹	2009	To explore the relationship between team behaviour and the delivery of an appropriate standard of care specific to the simulated case	Video observation of trainees participating in a simulated event involving the presentation of acute dyspnoea	Assessment of team behaviour using a behaviourally anchored rating scale (leadership and team behaviour measurement tool)	Checklist based expert rating
Catchpole <i>et al</i> ⁸²	2007	To investigate if effective teamwork can prevent the development of serious situations and provide evidence for improvements in training and systems	Live observation of surgical teams conducting paediatric cardiac and orthopaedic surgeries	Observation of non-technical skills using a behaviourally anchored rating scale (NOTECHS scoring system)	Assessment of minor problems, intraoperative performance and duration of surgery
Catchpole <i>et al</i> ⁸³	2008	To analyse the effects of surgical, anaesthetical and nursing teamwork skills on technical outcomes	Observation of surgical teams conducting laparoscopic cholecystectomies and carotid endarterectomies	Observation of non-technical skills using a behaviourally anchored rating scale (NOTECHS scoring system)	Operating time and errors in surgical technique

Continued

Table 1 Continued

Authors	Year	Main study objectives	Participants and setting	Teamwork process measure	Outcome measure
Cooper ⁸⁴	1999	To examine the relationship between leadership behaviour, team dynamics and task performance	Video observation of emergency teams managing full cardiopulmonary arrests with a resuscitation attempt lasting longer than 3 min	Survey about leadership behaviour using the Leadership Behaviour Description Questionnaire	Checklist based expert rating
Davenport et al ⁸⁵	2007	To measure the impact of organisational climate safety factors on risk-adjusted surgical morbidity and mortality	Survey of staff on general and vascular surgery services	Survey about teamwork climate, level of communication and collaboration with surgeon	Surgical morbidity Surgical mortality
El Bardissi et al ⁸⁶	2008	To identify patterns of teamwork failures that would benefit from intervention in the cardiac surgical setting	Live observation of surgical teams conducting cardiac surgery	Structured observation of teamwork failures that disrupted the flow of the operation	Surgical technical errors
Gillespie et al ⁸⁷	2012	To investigate how various human factors variables, extend the expected length of an operation	Live observation of surgical teams across 10 specialities	Structured observation of numbers of communication failures	Deviation from expected length of operation
Kolbe et al ⁸⁸	2012	To test the relationship between speaking up and technical team performance in anaesthesia.	Observation of two-person (nurse, resident) ad hoc anaesthesia teams performing simulated inductions of general anaesthesia with minor non-routine events	Structured observation of speaking up behaviour	Checklist based expert rating
Künzle et al ⁸⁹	2009	To investigate shared leadership patterns during anaesthesia induction and to show how they are linked to team performance	Observation of two-person (nurse, resident) ad hoc anaesthesia teams performing simulated inductions of general anaesthesia with a non-routine event (asystole)	Structured observation of leadership behaviour	Reaction time to non-routine event
Manojlovich et al ⁹⁰	2009	To determine the relationships between patients' outcomes and nurses' perceptions of communication and characteristics of the practice environment.	A survey was conducted with nurses on various ICU wards	Survey about perception of nurse-physician communication using the ICU-nurse physician questionnaire	Ventilator-associated pneumonia Bloodstream infections Pressure ulcers Acute physiology and chronic health evaluation score
Manser et al ⁹¹	2015	To investigate surgeons team management skills and its influence on performance	Live observation of surgical teams managing a simulated laparoscopic cholecystectomy	Structured observation of team management using the ComEd-E observation system	Checklist based expert rating
Marsch et al ⁹¹	2004	To determine whether and how human factors affect the quality of cardiopulmonary resuscitation	Observation of healthcare worker (nurse, physician) managing a cardiac arrest due to ventricular fibrillation using a high-fidelity patient simulator	Structured observation of task distribution, information transfer and leadership behaviour within the team	Checklist based expert rating

Continued

Table 1 Continued

Authors	Year	Main study objectives	Participants and setting	Teamwork process measure	Outcome measure
Mazzocco <i>et al</i> ⁹²	2009	To determine if patients of teams with good teamwork had better outcomes than those with poor teamwork	Live observation of surgical teams managing a variety of surgical procedures	Structured observation of information sharing, inquiry for relevant information and vigilance and awareness using a behaviourally anchored rating scale	Postoperative complications and death
Mishra <i>et al</i> ⁹³	2008	To report on the development and evaluation of a method for measuring operating-theatre teamwork quality	Live observation of surgical teams conducting laparoscopic cholecystectomy	Assessment of non-technical skills using a behaviourally anchored rating scale (NOTECHS scoring system)	Surgical technical errors assessed with the OCHRA-tool
Schmutz <i>et al</i> ⁹⁴	2015	To investigate the moderating effect of task characteristics on the relationship between coordination and performance	Video observation of paediatric teams managing various paediatric emergencies using a high-fidelity patient simulator	Structured observation of closed loop communication, task distribution and provide information without request using the CoMeT-E observation system	Checklist based expert rating
Siassakos <i>et al</i> ⁹⁵	2012	To investigate the relationship between patient satisfaction and communication	Video observation of teams (physicians, midwives) managing obstetrical emergencies in secondary and tertiary maternity units	Structured observation of closed loop communication	Timely administration of magnesium sulfate
Siassakos <i>et al</i> ⁹⁶	2011	To determine whether team performance in a simulated emergency is related to generic teamwork skills and behaviours	Video observation of healthcare professionals (physician, midwives) managing various emergencies using a high-fidelity patient simulator	Assessment of generic teamwork using a behaviourally anchored rating scale (teamwork analytical tool)	Clinical efficiency score
Thomas <i>et al</i> ⁹⁷	2006	To investigate the relationship of team behaviours during delivery room care and behaviours relate to the quality of care	Video observation of neonatal care teams managing a resuscitation during a caesarean section	Structured observation of communication, team management and leadership guidelines	Compliance with Neonatal Resuscitation Programme guidelines
Tschan <i>et al</i> ⁹⁸	2006	To investigate the influence of human factors on team performance in medical emergency driven groups	Video observation of medical emergency teams (senior doctor, resident, nurse) managing a cardiac arrest in a high-fidelity patient simulator	Structured observation of directive leadership and structuring inquiry	Clinical performance assessed based on a time-based coding of observable technical acts
Tschan <i>et al</i> ⁹⁹	2009	To investigate the influence of communication on diagnostic accuracy in ambiguous situations	Video observation of groups of physicians diagnosing a difficult patient with an anaphylactic shock in a high-fidelity patient simulator	Structured observation of the diagnostic information that have been considered, explicit reasoning and talking to the room	Accuracy of diagnosis

Continued

Table 1 Continued

Authors	Year	Main study objectives	Participants and setting	Teamwork process measure	Outcome measure
Westli <i>et al</i> ¹⁰⁰	2010	To investigate whether demonstrated teamwork skills and behaviour indicating shared mental models would be associated with improved medical management	Video observation of trauma teams (surgeons, anaesthesiologists, nurses, radiographers) in a high-fidelity patient simulator	Assessment of non-technical skills using a behaviourally anchored rating scale (ANTS and ATOM scoring system)	Checklist based expert rating
Wiegmann <i>et al</i> ¹⁰¹	2007	To investigate surgical errors and their relationship to surgical flow disruptions to understand better the effect of these disruptions on surgical errors and patient safety	Live observation of surgical teams conducting cardiac surgery operations	Structured observation of teamwork and communication failures	Structured observation of surgical errors during the operation
Williams <i>et al</i> ¹⁰²	2010	To describe relationships between teamwork behaviours and errors during neonatal resuscitation	Video observation of intensive care teams managing neonatal resuscitations	Structured observation of teamwork behaviour (vigilance, workload management, information sharing, inquiry, assertion)	Structured observation of errors (non-compliance with guidelines)
Wright <i>et al</i> ¹⁰³	2009	To test if observer ratings of team skills will correlate with objective measures of clinical performance	Video observation of teams consisting of medical students performing low-fidelity classroom based patient assessment and high-fidelity simulation emergent care	Observation using a behaviourally anchored rating scale for teamwork skills (assertiveness, decision-making, situation assessment, leadership, communication)	Checklist based expert rating
Yamada <i>et al</i> ¹⁰⁴	2016	To investigate the effect of standardised communication techniques on errors during resuscitation	Video observation of teams (Neonatologists, neonatal nurse practitioners, neonatology fellows) managing neonatal resuscitation	Structured observation of standardised communication	Error rate Time to initiate positive pressure ventilation Time to chest compression

ICU, intensive care unit.

Table 2 Studies, effect sizes and moderator variables included in the meta-analytical database

Authors	Year	Study goal	Setting	No. of teams	Professional composition	Team familiarity	Average team size	Task type	Patient realism	Performance measure
Amacher <i>et al</i> ⁶⁰	2017	0.11	Emergency medicine	72	Uniprofessional	Experi-mental	3	Non-routine	Simulated	Process
Brogaard <i>et al</i> ⁶⁰	2018	0.43	Obstetrics	99	Interprofessional	Real	5	Non-routine	Real	Process
Burtscher <i>et al</i> ⁸¹	2011	-0.27	Anaesthesia	31	Interprofessional	Experi-mental	2	Routine	Simulated	Process
Burtscher <i>et al</i> ⁸²	2011	0.19	Anaesthesia	15	Interprofessional	Experi-mental	2	Routine & non-routine	Simulated	Process
Burtscher <i>et al</i> ⁸³	2010	0.07	Anaesthesia	22	Interprofessional	Real	3	Non-routine	Real	Process
Carlson <i>et al</i> ⁹	2009	0.83	Emergency medicine	44	Uniprofessional	Experi-mental	2.6	Non-routine	Simulated	Process
Catchpole <i>et al</i> ⁸²	2007	.45†	Surgery	24	Interprofessional	Real	9	Non-routine	Real	Process
Catchpole <i>et al</i> ⁸²	2007	.29†	Surgery	18	Interprofessional	Real	5	Routine	Real	Process
Catchpole <i>et al</i> ⁸³	2008	.36†	Surgery	26	Interprofessional	Real		Routine	Real	Process
Catchpole <i>et al</i> ⁸³	2008	.09†	Surgery	22	Interprofessional	Real		Routine	Real	Process
Cooper ⁸⁴	1999	0.50	General care	20	Interprofessional	Real	4	Routine	Real	Process
Davenport <i>et al</i> ⁸⁵	2007	0.17	Surgery	52	Interprofessional	Real		Routine	Real	Outcome
El Bardissi <i>et al</i> ⁸⁶	2008	0.67	Surgery	31	Interprofessional	Real	7	Routine	Real	Process
Gillespie <i>et al</i> ⁸⁷	2012	0.23	Surgery	160	Interprofessional	Real	6	Routine	Real	Process
Kolbe <i>et al</i> ⁸⁸	2012	0.33	Anaesthesia	31	Interprofessional	Real	2	Non-routine	Simulated	Process
Künzle <i>et al</i> ⁸⁹	2009	0.56	Anaesthesia	12	Interprofessional	Real	2	Routine	Simulated	Process
Manojlovich <i>et al</i> ⁹⁰	2009	0.11	Intensive care	25	Uniprofessional	Real	36	Routine	Real	Outcome
Manser <i>et al</i> ⁶¹	2015	0.39	Surgery	19	Interprofessional	Experi-mental	5	Routine	Simulated	Process
Marsch <i>et al</i> ⁸¹	2004	0.23	Intensive care	16	Interprofessional	Experi-mental	3	Non-routine	Simulated	Process
Mazzocco <i>et al</i> ⁸²	2009	0.11	Surgery	293	Interprofessional	Real	6	Routine	Real	Outcome
Mishra <i>et al</i> ⁹³	2008	0.05	Surgery	26	Interprofessional	Real	6	Routine	Real	Process
Schmutz <i>et al</i> ⁸⁴	2015	0.12	Emergency medicine	68	Interprofessional	Real	6	Non-routine	Simulated	Process
Siassakos <i>et al</i> ⁸⁵	2012	0.66	Obstetrics	19	Interprofessional	Real	6	Non-routine	Simulated	Process
Siassakos <i>et al</i> ⁸⁶	2011	0.55	Emergency medicine/obstetrics	24	Interprofessional	Experi-mental	6	Non-routine	Simulated	Process
Thomas <i>et al</i> ⁸⁷	2006	0.23	Neonatal care	132	Interprofessional	Real	5	Non-routine	Real	Process

Continued

Table 2 Continued

Authors	Year	Study goal	Setting	No. of teams	Professional composition	Team familiarity	Average team size	Task type	Patient realism	Performance measure
Tschan <i>et al</i> ⁹⁸	2006	0.23	Emergency medicine	21	Interprofessional	Experimental	5	Non-routine	Simulated	Process
Tschan <i>et al</i> ⁹⁹	2009	0.37	Emergency medicine	20	Uniprofessional	Experimental	2.65	Non-routine	Simulated	Outcome
Westli <i>et al</i> ¹⁰⁰	2010	0.18	Emergency medicine	27	Interprofessional	Real	5.1	Non-routine	Simulated	Process
Wiegmann <i>et al</i> ¹⁰¹	2007	0.56	Surgery	31	Interprofessional	Real		Routine	Real	Process
Williams <i>et al</i> ¹⁰²	2010	0.18	Neonatal care	12	Interprofessional	Real	5	Non-routine	Real	Process
Wright <i>et al</i> ¹⁰³	2009	0.81	General care	9	Uniprofessional	Experimental	4	Non-routine	Simulated	Process
Yamada <i>et al</i> ¹⁰⁴	2016	0.11	Emergency medicine	13	Interprofessional	Experimental	3	Non-routine	Simulated	Process

*Carlson, Min & Bridges has been identified as an outlier and therefore excluded from the analysis.

†Effect sizes (r) with an † represent an average for a single sample and a single outcome and have been combined for this meta-analysis.

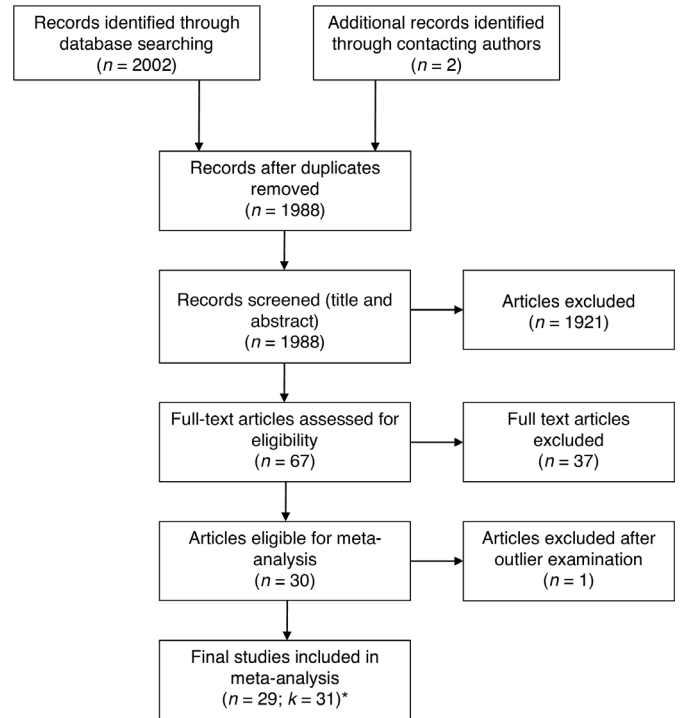


Figure 1 Systematic literature search.

and different forms of performance measures. Second, we aimed to provide an inference on the average effect in the entire population of studies from which the included studies are assumed to be a random selection of it. Therefore, random-effects models were estimated.⁵⁷ These models were calculated by the restricted maximum-likelihood estimator, an efficient and unbiased estimator.⁵⁸ Since we included only descriptive studies and no interventions we only included the sample size of the individual studies as a potential bias into the meta-analysis. To rule out a potential publication bias, we tested for funnel plot asymmetry using the random-effect version of the Egger test.⁵⁹ The results indicate that there is no asymmetry in the funnel plot ($z=1.79, p=0.074$), suggesting that there is no publication bias.

The estimation of meta-analytical models including the outlier analyses were performed with the package ‘metafor’ from the programming language and statistical environment R.⁵⁸

RESULTS

The online search resulted in 2002 articles (figure 1). Two studies were identified via contacting authors directly and have been presented at conferences in the past.^{60 61} After duplicates were removed 1988 articles were screened using title and abstract. Sixty-seven articles were then selected for a full text review. Full text examination, forward and backward search of selected articles and relevant reviews resulted in 30 studies coming from 28 articles (two publications presented two independent studies in one publication^{62 63}). This led to a total of 32 studies coming from 30 articles. Following the recommendation

by Viechtbauer and Cheung,⁶⁴ we screened for outliers using studentized deleted residuals. One case (Carlson *et al*,⁹ $r=0.89$, $n=44$, studentized deleted residuals=4.26) was identified as outlier and therefore excluded from further analyses, resulting in a final sample size of $k=31$.

Table 1 provides a qualitative description of the selected articles including study objectives, the setting in which the studies were carried out and a description of the teamwork processes as well as the outcome measures that were assessed. If a specific tool for the assessment of a teamwork process or outcome measure was used this is indicated in the corresponding column. Observational studies were most prevalent. Teamwork processes were assessed using either behaviourally anchored rating scales ($n=8$) or structured observation ($n=19$) of specific teamwork behaviour. Structured observation — as we describe it — is defined as a purely descriptive assessment of certain behaviour usually using a predefined observation system (eg, amount of speaking up behaviour). In contrast, behaviourally anchored rating scales consist of an evaluation of teamwork process behaviour by an expert. Only three studies used surveys to assess teamwork behaviours. The majority of the studies ($n=27$) assessed process performance using either a checklist based

expert rating or assessing a reaction time measure after the occurrence of a certain event (eg, time until intervention). Only four studies assessed outcome performance measures. Measures included accuracy of diagnosis, post-operative complications and death, surgical morbidity and mortality, ventilator-associated pneumonia, bloodstream infections, pressure ulcers and acute physiology and chronic health evaluation score. Table 2 provides an overview of all variables included in the meta-analysis including the effect sizes and moderator variables.

Effect of teamwork and contextual factors

Table 3 and figure 2 shows the relationship between teamwork and team performance. The sample-sized weighted mean correlation was 0.28 (95% CI 0.20 to 0.35, $z=6.55$, $p<0.001$), indicating that teamwork is positively related to clinical performance. Results further indicated heterogeneous effect size distributions across the included samples ($Q=53.73$, $p<0.05$, $I^2=45.96$), signifying that the variability across the sample effect sizes was more than what would be expected from sampling error alone.

To test for moderator effects of the contextual factors, we conducted mixed-effects models including the mentioned moderators: *professional composition*, *team*

Table 3 Meta-analytical relationships between teamwork and clinical performance

	N	k	r	95% CI	80% CR	Q	I ²
Overall relationship	1390	31	0.28*	(0.20 to 0.35)	(0.09 to 0.45)	53.7*	46.0
Team characteristics							
Professional composition							
Interprofessional	1264	27	0.28*	(0.20 to 0.36)	(0.09 to 0.46)	47.1*	48.2
Uniprofessional	126	4	0.28	(−0.01 to 0.52)	(−0.04 to 0.54)	6.5	47.1
Team familiarity							
Experimental team	240	10	0.25*	(0.05 to 0.43)	(−0.05 to 0.51)	17.2*	47.2
Real team	1150	21	0.29*	(0.20 to 0.37)	(0.12 to 0.45)	36.2*	45.7
Team size†							
Task characteristics							
Task type							
Routine task	766	14	0.27*	(0.12 to 0.40)	(−0.01 to 0.50)	30.9*	65.0
Non-routine task	609	16	0.29*	(0.20 to 0.39)	(0.16 to 0.42)	20.5	24.6
Methodological factors							
Patient realism							
Real patient	993	16	0.28*	(0.18 to 0.38)	(0.10 to 0.45)	28.7*	49.3
Simulated patient	397	15	0.28*	(0.13 to 0.41)	(0.02 to 0.50)	25.0*	44.6
Performance measures							
Outcome performance	390	4	0.13*	(0.03 to 0.23)	(0.06 to 0.19)	1.3	0.0
Process performance	1000	27	0.30*	(0.21 to 0.39)	(0.10 to 0.49)	45.6*	45.6

* $p < .05$.

I² = % of total variability in the effect size estimates due to heterogeneity among true effects (vs sampling error).

†Team size was entered as a continuous variable, therefore, no subgroup analyses exist.

CI, confidence interval; CR, credibility interval; K, number of studies; N, cumulative sample size (number of teams); Q, test statistic for residual heterogeneity of the models; r, sample-size weighted correlation.

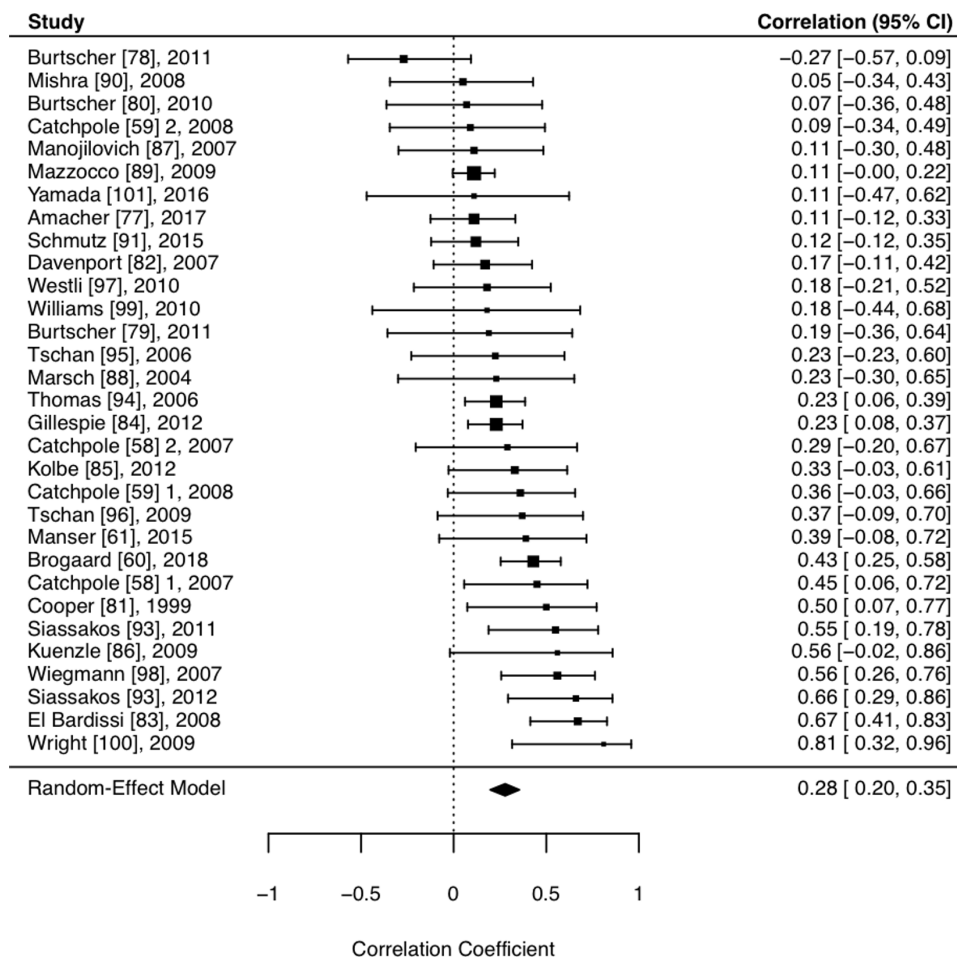


Figure 2 Relationship between teamwork processes and performance.

familiarity, team size, task type, patient realism and performance measures.

The omnibus test of moderators was not significant ($F=0.18$, $df_1=6$, $df_2=18$, $p>0.20$), suggesting that the examined contextual factors did not influence the average effect of teamwork on clinical performance. To provide greater detail about the role of the contextual factors, we conducted separate analyses for the categorical contextual factors and report them in [table 3](#).

DISCUSSION

With this study, we aimed to provide evidence for the performance implications of teamwork in healthcare teams. By including various contextual factors, we investigated potential contingencies that these factors might have on the relationship between teamwork and clinical performance. The analysis of 1390 teams from 31 different studies showed that teamwork has a medium sized effect ($r=0.28^{65,66}$;) on clinical performance across various care settings. Our study is the first to investigate this relationship quantitatively with a meta-analytical procedure. This finding aligns with and advances previous work that explored this relationship in a qualitative way.^{8 15 17 43–47}

At first glance a correlation of $r=0.28$ might not seem very high. However, we would like to highlight that $r=0.28$ is considered a medium sized effect^{65 66} and should not be underestimated. To better illustrate what this effect means we transformed the correlation into an OR of 2.8.⁵³ Of course, this transformation simplifies the correlation because teamwork and often the outcome measures are not simple dichotomous variables that can be divided into an intervention and control group. However, this transformation illustrates that teams who engage in teamwork processes are 2.8 times more likely to achieve high performance than teams who are not. Looking at the performance measures in our study we see that they either describe patient outcomes (eg, mortality, morbidity) or are closely related to patient outcomes (eg, adherence to treatment guidelines). Thus, we consider teamwork a performance-relevant process that needs to be promoted through training and implementation into treatment guidelines and policies.

The included studies used a variety of different measures for clinical performance. This variability resulted from the different clinical contexts in which the studies were carried out. There is no universal measure for clinical performance because the outcome is in most cases context specific. In surgery, common

performance measures are surgical complications, mortality or morbidity.⁶⁷ In anaesthesia, studies often use expert ratings based on checklists to assess the provision of anaesthesia. Expert ratings are also the common form of performance assessment in simulator settings where patient outcomes like morbidity or mortality cannot be measured. Future studies need to be aware that clinical performance measures depend on the clinical context and that the development of valid performance measures requires considerable effort and scientific rigour. Guidelines on how to develop performance assessment tools for specific clinical scenarios exist and need to be accounted for.^{52 68 69} Furthermore, depending on the clinical setting researchers need to evaluate what specific clinical performance measures are suitable and if and how they can be linked to team processes in a meaningful way.

The analysis of moderators illustrates that teamwork is related with performance under a variety of conditions. Our results suggest that teams in different contexts characterised by different team constellations, team size and levels of acuity of care all benefit from teamwork. Therefore, clinicians and educators from all fields should strive to maintain or increase effective teamwork. In recent years, there has been an upsurge in crisis resource management (CRM).¹⁹ These trainings focus on team management and implement various teamwork principles during crisis situations (eg, emergencies).⁷⁰ Our results suggest that team trainings should not only focus on non-routine situations like emergencies but also on routine situations (eg, routine anaesthesia induction, routine surgery) because based on our data teamwork is equally important in such situations.

A closer look at methodological factors of the included studies revealed that the observed relationship between teamwork and performance in simulation settings does not differ from relationships observed in real settings. Therefore, we conclude that teamwork studies conducted in simulation settings generalise to real life settings in acute care. Further, the analysis of different performance measures reveals a trend towards process performance measures being more strongly related with teamwork than outcome performance measures. A possible explanation of this finding relates to the difficulty of investigating outcome performance measures in a manner isolated from other variables. Nevertheless, we still found a significant relationship between teamwork and objective patient outcomes (eg, postoperative complications, bloodstream infections) despite the methodological challenges of measuring outcome performance and the small number of studies using outcome performance ($k=4$).

Our results are in line with previous meta-analyses investigating the effectiveness of team training in healthcare.^{18 19} Similar to our results, Hughes *et al* highlighted the effectiveness of team trainings under a variety of conditions — irrespective of team composition,¹⁸

simulator fidelity or patient acuity of the trainee's unit as well as other factors.

We were unable to find a moderation of task type in our study, potentially explained by task interdependence, which reflects the degree to which team members depend on one another for their effort, information and resources.⁷¹ A meta-analysis including teams from multiple industries (eg, project teams, management teams) found that task interdependence moderates the relationship between teamwork and performance, demonstrating the importance of teamwork for highly interdependent team tasks.⁷² Most studies included in our analysis focused on rather short and intense patient care episodes (eg, a surgery, a resuscitation task) with high task interdependence, which may explain the high relevance of teamwork for all these teams.

Limitations and future directions

Despite greater attention to healthcare team research and team training over the last decade, we were only able to identify 32 studies (31 included in the meta-analysis). Of note, over two-thirds of the studies in our analysis emerged in the last 10 years, reflecting the increasing interest in the topic. The rather small number of studies might relate to the difficulties in quantifying teamwork, the considerable theoretical and methodological knowledge required and the challenges of capturing relevant outcome measures. Also, besides the manual searches of selected articles and reviews and contacting authors in the field we did only search the database PubMed. PubMed is the most common database to access papers that potentially investigate medical teams and includes approximately 30 000 journals from the field of medicine, psychology and management. We are fairly confident that through the additional inclusion of relevant reviews and forward and backwards search, our results represent an accurate representation of what can be found in the literature.

Future research should build on recent theoretical and applied work^{24 26 28 73} about teamwork and use this current meta-analysis as a signpost for future investigations. In order to move our field forward, we must use existing conceptual frameworks^{22 24 26} and establish standards for investigating teams and teamwork. This can often only be achieved with interdisciplinary research teams including experts from the medical fields but equally important from health professions education, psychology or communication studies.

Another limitation relates to the unbalanced analysis of subgroups. For example, we only identified four studies that used outcome performance variables compared with 27 using process performance measures. Uneven groups may reduce the power to detect significant differences. Therefore, we encourage future studies to include outcome performance measures despite the effort required.

Finally, more factors may influence the relationship between teamwork and performance that we

were unable to extract from the studies. While we tested for the effects of team familiarity by comparing experimental teams and real teams, this does not fully capture team member familiarity. The extent to which team members actually worked together during prior clinical practice might predict of how effectively they perform together. However, even two people working in the same ward might actually not have interacted much during patient care depending on the setting. Also team climate on a ward or in a hospital may be an important predictor of how well teams work together, especially related to sharing information or speaking up within the team.^{74 75}

Finally, the clinical context might play a role in how team members collaborate. In different disciplines, departments or healthcare institutions different norms and routines exist on how to work together. Therefore, study results and recommendations about teamwork need to be interpreted in the light of the respective clinical context. There are empirical indications that a one-size-fits-all approach might not be suitable and team training efforts cannot ignore the clinical context, especially the routines and norms about collaboration.⁷⁶ We acknowledge that there might be other factors surrounding healthcare teams that might potentially influence teamwork and clinical performance. However, in this review we could only extract data that was reported in the primary studies. Since these were limited in the healthcare contexts studied, the results might not generalise to long-term care settings or mental health, for example. Future work needs to consider and also document a broader range of potentially influencing factors.

CONCLUSION

The current meta-analysis confirms that teamwork across various team compositions represents a powerful process to improve patient care. Good teamwork can be achieved by joint reflection about teamwork during clinical event debriefings^{77 78} as well as team trainings⁷⁹ and system improvement. All healthcare organisations should recognise these findings and place continuous efforts into maintaining and improving teamwork for the benefit of their patients.

Acknowledgements The authors thank Manuel Stühlinger for his help with study selection and data extraction and Walter J. Eppich, MD, PhD for a critical review and proofreading the manuscript.

Contributors All authors substantially contributed to this study and were involved in the study design. JS drafted the paper. LM analysed the data and revised the manuscript for content. TM revised the manuscript for content and language. All authors approved the final version.

Funding This work was funded by the European Society of Anaesthesiology (ESA) and the Swiss National Science Foundation (SNSF, Grant No. P300P1_177695). The ESA provided resources for an additional research assistant helping with literature search and selection. Part of the salary of JS was funded by the SNSF.

Competing interests None declared.

Patient and public involvement statement Patients and public were not involved in this study.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Thomas EJ. Improving teamwork in healthcare: current approaches and the path forward. *BMJ Qual Saf* 2011;20:647–50.
2. Salas E, Rosen MA. Building high reliability teams: progress and some reflections on teamwork training. *BMJ Qual Saf* 2013;22:369–73.
3. Schraagen JM, Schouten T, Smit M, *et al*. A prospective study of paediatric cardiac surgical microsystems: assessing the relationships between non-routine events, teamwork and patient outcomes. *BMJ Qual Saf* 2011;20:599–603.
4. Williams AL, Lasky RE, Dannemiller JL, *et al*. Teamwork behaviours and errors during neonatal resuscitation. *Quality and Safety in Health Care* 2010;19:60–4.
5. Guerlain S, Turrentine FE, Bauer DT, *et al*. Crew resource management training for surgeons: feasibility and impact. *Cognition, Technology & Work* 2008;10:255–64.
6. McCulloch P, Mishra A, Handa A, *et al*. The effects of aviation-style non-technical skills training on technical performance and outcome in the operating theatre. *Qual Saf Health Care* 2009;18:109–15.
7. Yule S, Flin R, Maran N, *et al*. Debriefing surgeons on non-technical skills (NOTSS). *Cogn Technol Work* 2007;139:131–274.
8. Schmutz J, Manser T. Do team processes really have an effect on clinical performance? A systematic literature review. *Br J Anaesth* 2013;110:529–44.
9. Carlson J, Min E, Bridges D. The impact of leadership and team behavior on standard of care delivered during human patient simulation: a pilot study for undergraduate medical students. *Teach Learn Med* 2009;21:24–32.
10. Mishra A, Catchpole K, McCulloch P. The Oxford NOTECHS system: reliability and validity of a tool for measuring teamwork behaviour in the operating theatre. *Qual Saf Health Care* 2009;18:104–8.
11. Manojlovich M, DeCicco B. Healthy work environments, nurse-physician communication, and patients' outcomes. *Am J Crit Care* 2007;16:536–43.
12. Schmutz J, Hoffmann F, Heimberg E, *et al*. Effective coordination in medical emergency teams: the moderating role of task type. *Eur J Work Organ Psychol* 2015;24:761–76.
13. Tschan F, Semmer NK, Nägele C, *et al*. Task adaptive behavior and performance in groups. *Group Processes & Intergroup Relations* 2000;3:367–86.
14. Manser T. Teamwork and patient safety in dynamic domains of healthcare: a review of the literature. *Acta Anaesthesiol Scand* 2009;53:143–51.
15. Fernandez Castela E, Russo SG, Riethmüller M, *et al*. Effects of team coordination during cardiopulmonary resuscitation: a systematic review of the literature. *J Crit Care* 2013;28:504–21.
16. Dietz AS, Pronovost PJ, Benson KN, *et al*. A systematic review of behavioural marker systems in healthcare: what do we know about their attributes, validity and application? *BMJ Qual Saf* 2014;23:1031–9.
17. Flowerdew L, Brown R, Vincent C, *et al*. Identifying nontechnical skills associated with safety in the emergency department: a scoping review of the literature. *Ann Emerg Med* 2012;59:386–94.
18. Hughes AM, Gregory ME, Joseph DL, *et al*. Saving lives: a meta-analysis of team training in healthcare. *J Appl Psychol* 2016;101:1266–304.
19. O'Dea A, O'Connor P, Keogh I. A meta-analysis of the effectiveness of Crew resource management training in acute care domains. *Postgrad Med J* 2014;90:699–708.
20. Salas E, Rosen MA, King H. Managing teams managing crises: principles of teamwork to improve patient safety in the emergency room and beyond. *Theor Issues Ergon Sci* 2007;8:381–94.

21. Salas E, Stagl KC, Burke CS. *Fostering team effectiveness in organizations: toward an integrative theoretical framework*, 2007: 52.
22. Fernandez R, Kozlowski SWJ, Shapiro MJ, et al. Toward a definition of teamwork in emergency medicine. *Acad Emerg Med* 2008;15:1104–12.
23. Bowers CA, Braun CC, Morgan BB. Team workload: Its meaning and measurement. In: Brannik MT, Salas E, Prince C, eds. *Team performance assessment and measurement: theory, methods, and applications*. New York: Team performance assessment and measurement: Theory, methods, and applications, 1997: 85–108.
24. Marks MA, Mathieu JE, Zaccaro SJ. A temporally based framework and taxonomy of team processes. *Acad Manage Rev* 2001;26:356–76.
25. Dietz AS, Pronovost PJ, Mendez-Tellez PA, et al. A systematic review of teamwork in the intensive care unit: what do we know about teamwork, team tasks, and improvement strategies? *J Crit Care* 2014;29:908–14.
26. Reader TW, Flin R, Mearns K, et al. Developing a team performance framework for the intensive care unit. *Crit Care Med* 2009;37:1787–93.
27. Ilgen DR, Hollenbeck JR, Johnson M, et al. Teams in organizations: from input-process-output models to IMOI models. *Annu Rev Psychol* 2005;56:517–43.
28. Burke CS, Stagl KC, Salas E, et al. Understanding team adaptation: a conceptual analysis and model. *J Appl Psychol* 2006;91:1189–207.
29. Lemieux-Charles L, McGuire WL. What do we know about health care team effectiveness? A review of the literature. *Med Care Res Rev* 2006;63:263–300.
30. Kozlowski SWJ. Advancing research on team process dynamics: theoretical, methodological, and measurement considerations. *Organizational Psychology Review* 2015;5:270–99.
31. Maloney MM, Bresman H, Zellmer-Bruhn ME, et al. Contextualization and context theorizing in teams research: a look back and a path forward. *Academy of Management Annals* 2016;10:891–942.
32. Kozlowski SW, Klein KJ. *A multilevel approach to theory and research in organizations: contextual, temporal, and emergent processes*, 2000.
33. Schofield RF, Amodeo M. Interdisciplinary teams in health care and human services settings: are they effective? *Health Soc Work* 1999;24:210–9.
34. Hall P. Interprofessional teamwork: professional cultures as barriers. *J Interprof Care* 2005;19 Suppl 1:188–96.
35. Lillrank P. The quality of standard, routine and nonroutine processes. *Organization Studies* 2003;24:215–33.
36. Eppich WJ, O'Connor L, Adler M. Providing effective simulation activities. In: Forrest K, McKimm J, Edgar S, eds. *Essential simulation in clinical education*. Oxford, UK: John Wiley & Sons, Ltd, 2013: 213–34.
37. Anderson N, Ones DS, Sinangil HK, et al. *Handbook of industrial, work and organizational psychology: personnel psychology*. London: Sage Publications Ltd, 2001.
38. Sonnentag S, Frese M. Performance concepts and performance theory. In: *Psychological management of individual performance*. , 2002: 23, 3–25.
39. Cooper S, Cade J. Predicting survival, in-hospital cardiac arrests: resuscitation survival variables and training effectiveness. *Resuscitation* 1997;35:17–22.
40. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;6:e1000100.
41. Hunter JE, Schmidt FLOaks T, ed. *Methods of meta-analysis: correcting error and bias in research findings*. CA: Sage Publications, 2004.
42. Card N. *Applied meta-analysis for social science research*. New York: Guilford Publications, 2015.
43. Fletcher GCL, McGeorge P, Flin RH, et al. The role of non-technical skills in anaesthesia: a review of current literature. *Br J Anaesth* 2002;88:418–29.
44. Ghaferi AA, Dimick JB, teamwork lof. Importance of teamwork, communication and culture on failure-to-rescue in the elderly. *Br J Surg* 2016;103:e47–51.
45. Mitchell L, Flin R. Non-Technical skills of the operating theatre scrub nurse: literature review. *J Adv Nurs* 2008;63:15–24.
46. Santos R, Bakero L, Franco P, et al. Characterization of non-technical skills in paediatric cardiac surgery: communication patterns. *Eur J Cardiothorac Surg* 2012;41:1005–12.
47. Youngson GG. Nontechnical skills in pediatric surgery: factors influencing operative performance. *J Pediatr Surg* 2016;51:226–30.
48. Kianfar S, Carayon P, Hundt AS, et al. Care coordination for chronically ill patients: identifying coordination activities and interdependencies. *Appl Ergon* 2019;80:9–16.
49. Gully SM, Devine DJ, Whitney DJ. A meta-analysis of cohesion and performance effects of level of analysis and task interdependence. *Small Group Research* 1995;26:497–520.
50. Hunter JE, Schmidt FL. Dichotomization of continuous variables: the implications for meta-analysis. *J Appl Psychol* 1990;75:334–49.
51. Campbell JP. Modeling the performance prediction problem in industrial and organizational psychology. In: Dunnette MD, Hough LM, eds. *Handbook of industrial and organizational psychology*. Palo Alto: CA: Consulting Psychologists Press, 1990: 687–732.
52. Schmutz J, Eppich WJ, Hoffmann F, et al. Five steps to develop checklists for evaluating clinical performance: an integrative approach. *Acad Med* 2014;89:996–1005.
53. Borenstein M, Hedges LV, Higgins JPT, et al. *Introduction to meta-analysis*. Hoboken, NJ: John Wiley & Sons, Ltd, 2009.
54. Walker DA. JMASM9: converting Kendall's tau for correlational or meta-analytic analyses. *J Mod App Stat Meth* 2003;2:525–30.
55. Lipsey MW, Wilson D. *Practical meta-analysis*. New York: Sage Publications, Inc, 2001.
56. Rambo WW, Chomiak AM, Price JM. Consistency of performance under stable conditions of work. *J Appl Psychol* 1983;68:78–87.
57. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. *J Stat Softw* 2010;36:1–48.
58. Viechtbauer W. Bias and efficiency of meta-analytic variance estimators in the random-effects model. *J Educ Behav Stat* 2005;30:261–93.
59. Sterne JA, Egger M. Regression methods to detect publication and other bias in meta-analysis. In: *Publication bias in meta-analysis: prevention, assessment and adjustments*. Chichester: John Wiley & Sons, 2006: 99–110.
60. Brogaard L, Kierkegaard O, Hvidmand L. Is non-technical performance the key to high clinical performance in obstetric teams? *th Annual conference of the British Maternal Fetal Medicine Society*, 2018.
61. Manser T, Bogdanovic J, Clack L, et al. Surgeon's team management skills predict clinical performance. In: *Paper presented at the annual meeting of the society for simulation applied to Medicine (SESAM), Lisbon*. Portugal, 2015.
62. Catchpole KR, Giddings AEB, Wilkinson M, et al. Improving patient safety by identifying latent failures in successful operations. *Surgery* 2007;142:102–10.
63. Catchpole K, Mishra A, Handa A, et al. Teamwork and error in the operating room. *Ann Surg* 2008;247:699–706.
64. Viechtbauer W, Cheung MW-L. Outlier and influence diagnostics for meta-analysis. *Res Synth Methods* 2010;1:112–25.
65. Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum, 1988.
66. Bosco FA, Aguinis H, Singh K, et al. Correlational effect size benchmarks. *J Appl Psychol* 2015;100:431–49.
67. Haynes AB, Weiser TG, Berry WR, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med Overseas Ed* 2009;360:491–9.
68. Gaba DM, Howard SK, Flanagan B, et al. Assessment of clinical performance during simulated crises using both technical and behavioral ratings. *Anesthesiology* 1998;89:8–18.
69. Rosen MA, Pronovost PJ. Advancing the use of checklists for evaluating performance in health care. *Academic Medicine* 2014;89:963–5.
70. Gaba DM, Howard SK, Fish KJ, et al. Simulation-based training in anesthesia crisis resource management (ACRM): a decade of experience. *Simul Gaming* 2001;32:175–93.
71. Wageman R, Baker G. Incentives and cooperation: the joint effects of task and reward interdependence on group performance. *J Organ Behav* 1997;18:139–58.
72. LePINE JA, Piccolo RF, Jackson CL, et al. A meta-analysis of teamwork processes: tests of a multidimensional model and relationships with team effectiveness criteria. *Pers Psychol* 2008;61:273–307.
73. Maynard MT, Kennedy DM, Sommer SA. Team adaptation: A fifteen-year synthesis (1998–2013) and framework for how this literature needs to “adapt” going forward. *Eur J Work Organ Psychol* 2015;24:652–77.
74. Edmondson AC. *Teaming: how organizations learn, innovate, and compete in the knowledge economy*. San Francisco, CA: Jossey-Bass, 2012.
75. Edmondson AC, Lei Z. Psychological safety: the history, renaissance, and future of an interpersonal construct. *Annu Rev Organ Psychol Organ Behav* 2014;1:23–43.

76. Henrickson Parker S, Schmutz JB, Manser T. Training needs for adaptive coordination: utilizing task analysis to identify coordination requirements in three different clinical settings. *Group Organ Manag* 2018;43:504–27.
77. Eppich WJ, Mullan PC, Brett-Fleegler M, et al. “Let’s Talk About It”: Translating Lessons From Health Care Simulation to Clinical Event Debriefings and Coaching Conversations. *Clin Pediatr Emerg Med* 2016;17:200–11.
78. Schmutz JB, Eppich WJ. Promoting learning and patient care through shared reflection: a conceptual framework for team reflexivity in health care. *Acad Med* 2017;92:1555–63.
79. Salas E, DiazGranados D, Weaver SJ, et al. Does team training work? principles for health care. *Acad Emerg Med* 2008;15:1002–9.
80. Amacher SA, Schumacher C, Legeret C, et al. Influence of gender on the performance of cardiopulmonary rescue teams: a randomized, prospective simulator study. *Crit Care Med* 2017.
81. Burtscher MJ, Kolbe M, Wacker J, et al. Interactions of team mental models and monitoring behaviors predict team performance in simulated anesthesia inductions. *J Exp Psychol Appl* 2011;17:257–69.
82. Burtscher MJ, Manser T, Kolbe M, et al. Adaptation in anaesthesia team coordination in response to a simulated critical event and its relationship to clinical performance. *Br J Anaesth* 2011;106:801–6.
83. Burtscher MJ, Wacker J, Grote G, et al. Managing nonroutine events in anesthesia: the role of adaptive coordination. *Hum Factors* 2010;52:282–94.
84. Cooper S, Wakelam A. Leadership of resuscitation teams: “Lighthouse Leadership”. *Resuscitation* 1999;42:27–45.
85. Davenport DL, Henderson WG, Mosca CL, et al. Risk-Adjusted morbidity in teaching hospitals correlates with reported levels of communication and collaboration on surgical teams but not with scale measures of teamwork climate, safety climate, or working conditions. *J Am Coll Surg* 2007;205:778–84.
86. ElBardissi AW, Wiegmann DA, Henrickson S, et al. Identifying methods to improve heart surgery: an operative approach and strategy for implementation on an organizational level. *Eur J Cardiothorac Surg* 2008;34:1027–33.
87. Gillespie BM, Chaboyer W, Fairweather N. Factors that influence the expected length of operation: results of a prospective study. *BMJ Qual Saf* 2012;21:3–12.
88. Kolbe M, Burtscher MJ, Wacker J, et al. Speaking up is related to better team performance in simulated anesthesia inductions: an observational study. *Anesth Analg* 2012;115:1099–108.
89. Künzle B, Zala-Mezo E, Wacker J, et al. Leadership in anaesthesia teams: the most effective leadership is shared. *BMJ Qual Saf* 2010;19:e46–6.
90. Manojlovich M, Antonakos CL, Ronis DL. Intensive care units, communication between nurses and physicians, and patients’ outcomes. *Am J Crit Care* 2009;18:21–30.
91. Marsch SCU, Müller C, Marquardt K, et al. Human factors affect the quality of cardiopulmonary resuscitation in simulated cardiac arrests. *Resuscitation* 2004;60:51–6.
92. Mazzocco K, Petitti DB, Fong KT, et al. Surgical team behaviors and patient outcomes. *Am J Surg* 2009;197:678–85.
93. Mishra A, Catchpole K, Dale T, et al. The influence of non-technical performance on technical outcome in laparoscopic cholecystectomy. *Surg Endosc* 2008;22:68–73.
94. Schmutz J, Hoffmann F, Heimberg E, et al. Effective coordination in medical emergency teams: the moderating role of task type. *Europ J Work Org Psych* 2015;24:761–76.
95. Siassakos D, Bristowe K, Draycott TJ, et al. Clinical efficiency in a simulated emergency and relationship to team behaviors: a multisite cross-sectional study. *Obstet Anesth Digest* 2012;32.
96. Siassakos D, Fox R, Crofts JF, et al. The management of a simulated emergency: better teamwork, better performance. *Resuscitation* 2011;82:203–6.
97. Thomas EJ, Sexton JB, Lasky RE, et al. Teamwork and quality during neonatal care in the delivery room. *J Perinatol* 2006;26:163–9.
98. Tschan F, Semmer NK, Gautschi D, et al. Leading to recovery: group performance and coordinative activities in medical emergency driven groups. *Hum Perform* 2006;19:277–304.
99. Tschan F, Semmer NK, Gurtner A, et al. Explicit Reasoning, confirmation bias, and illusory transactive memory: a simulation study of group medical decision making. *Small Group Res* 2009;40:271–300.
100. Westli HK, Johnsen BH, Eid J, et al. Teamwork skills, shared mental models, and performance in simulated trauma teams: an independent group design. *Scand J Trauma Resusc Emerg Med* 2010;18:47.
101. Wiegmann DA, ElBardissi AW, Dearani JA, et al. Disruptions in surgical flow and their relationship to surgical errors: an exploratory investigation. *Surgery* 2007;142:658–65.
102. Williams AL, Lasky RE, Dannemiller JL, et al. Teamwork behaviours and errors during neonatal resuscitation. *Qual Saf Health Care* 2010;19:60–4.
103. Wright MC, Phillips-Bute BG, Petrusa ER, et al. Assessing teamwork in medical education and practice: relating behavioural teamwork ratings and clinical performance. *Med Teach* 2009;31:30–8.
104. Yamada NK, Fuerch JH, Halamek LP. Impact of standardized communication techniques on errors during simulated neonatal resuscitation. *Am J Perinatol* 2016;33:385–92.