BMJ Open

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

| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2018-028280 |
| Article Type: | Research |
| Date Submitted by the Author: | 29-Nov-2018 |
| Complete List of Authors: | Schmutz, Jan B.; Northwestern University, Department of Communication Studies<br>Meier, Laurenz ; University of Neuchâtel , Occupational Psychology and Organizations<br>Manser, Tanja; University of Applied Sciences and Arts Northwestern Switzerland |
| Keywords: | teamwork, non-technical skills, communication, meta-analysis, teams |
| | |

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

**Jan B. Schmutz, PhD**

Researcher at the Department of Communication Studies, Northwestern University, Evanston, USA

Postdoctoral Researcher and Lecturer at the Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland

jan.schmutz@northwestern.edu

**Laurenz L. Meier, PhD**

Professor at the Institute of Work and Organizational Psychology, University of Neuchâtel, Neuchâtel, Switzerland

laurenz.meier@unine.ch

**Tanja Manser, PhD**

Director of the FHNW School of Applied Psychology, University of Applied Sciences and Arts

Northwestern Switzerland, Olten, Switzerland

tanja.manser@fhnw.ch



Word count: 3807

Keywords: teamwork, non-technical skills, communication, teams, meta-analysis

**Correspondence to**

Jan B. Schmutz
ATLAS Lab
Department of Communication Studies
Northwestern University
Frances Searle Building 1.220
2240 Campus Drive
Evanston, IL 60208
United States
Phone 847-491-7023
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.

**Methods** Data sources were searched up to June 2018 and included PubMed, manual backward search of relevant reviews, manual backward and forward search of studies included in the meta-analysis and contacting of selected authors via e-mail. Studies were included if they reported a relationship between a teamwork process and a performance measure. Moderator variables (i.e. professional composition, team familiarity, average teams size, task type, patient realism and type of performance measure) were coded and random-effect models were estimated. Two investigator independently extracted information on study characteristics in accordance with PRISMA 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 $r$ = .28, 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 organizations should recognize the value of teamwork and emphasize approaches that maintain and improve teamwork for the benefit of their patients.

## ARTICLE SUMMARY

### Strengths and limitations of this study

- This systematic review evaluates available studies investigating the effectiveness of teamwork processes.
- 31 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.

## INTRODUCTION

Teams are ubiquitous in healthcare and must work across professional, disciplinary and sectorial boundaries. Experts agree that effective teamwork anchors safe and effective care at various levels of the healthcare systems[1-4] leading to a relatively recent shift towards team research and training.[5-7] Our field now widely accepts that a team of individual experts does not necessarily make an expert team.[8,9]

However, the literature investigating healthcare teams reports mixed and sometimes even contradicting results about the relationship between teamwork and clinical performance.[9] Some studies find a large effect of teamwork on performance (e.g. Carlson et al.[10]) while others report small or no relationships.[11,12] 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 organization 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 (e.g. team composition and size, task characteristics, team environment) that likely influence the effect that teamwork has on clinical performance.[13,14]

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[9,15-18] 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[19,20] focused mainly on the effectiveness of team trainings but not on the effect of teamwork itself. This meta-analysis will generate strong quantitative evidence to inform the relevance of future interventions targeting teamwork in healthcare organizations.

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

**Teamwork and performance**

*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 behaviors. *Teamwork* is a process that describes interactions among team members who combine collective resources to resolve task demands (e.g. giving clear orders).[21,22] Teamwork or team processes can be differentiated from *taskwork*. *Taskwork* denotes a team's individual interaction with tasks, tools, machines and systems.[22] *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.[23] Therefore, *team performance* represents the accumulation of teamwork and taskwork (i.e. what the team actually does).[17]

Team performance is often described in terms of inputs, processes and outputs (IPO).[21,24-26] *Outputs* like quality of care, errors or performance are influenced by team related *processes* (i.e. 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. This IPO framework helps to systematize the mechanisms that predict team performance and represents the basis for the selection of the teamwork studies included in our meta-analysis.

**Contextual factors of teamwork effectiveness**

Based on a large body of team research from various domains, we hypothesize that several contextual and methodological factors might moderate the effectiveness of teamwork, indicating that teamwork is more important under certain conditions.[27,28] Therefore, we investigate several factors: (a) team characteristics (i.e. professional composition, team familiarity, team size); (b) task type (i.e. routine vs. non-routine tasks); (c) two methodological factors related to patient realism (i.e. simulated vs. real) and the type of performance measures used (i.e. 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.[29] Diverse educational paths in interprofessional teams may shape respective values, beliefs, attitudes and behaviors.[30] 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 to 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 to smaller teams.

*Task type.* Routine situations are characterized by repetitive and unvarying actions (e.g. standard anaesthesia induction).[31] 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.[32] 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 standardized 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 simulated patients 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 generalize 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- and outcome-related aspects of performance.[33,34] Process performance measures are action-related aspects and refer to adequate behaviour during procedures (e.g. adhering to guidelines), making them easier to assess. Outcome performance measures (e.g. 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 (e.g. 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 (i.e. performance outcome) of a cardiac arrest patient are *"duration of the arrest"* and *"age of the patient less than 70"*.[35] Although a team delivers perfect basic life support (i.e. high process performance) the patient can still die (i.e. 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 PRISMA statement[36] as well as established guidelines in social sciences.[37,38] 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 data base PubMed (no limit was placed on date of publication, last search June 2018) using the key words *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 8 systematic literature reviews that focus on teamwork or non-technical skills in various healthcare domains ,[9,16,18,39-43] c) a manual backwards search for all references cited in studies we included in our meta-analysis, d) a manual forward search to identify studies that cite the studies we included in our meta-analysis, e) identification of relevant unpublished manuscripts via e-mail 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 (e.g. coordination, communication). In addition, studies needed to investigate the relationship between at least one teamwork process and a performance measure (e.g. patient outcome). When studies reported multiple estimates of the same relationship from the same sample (e.g. 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.[44] All articles included in this meta-analysis are listed in Table 1 and Table 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-analytic integrations.[45,46]

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. 20% of the studies were rated by both coders. Intercoder 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 (e.g. 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 behavior often under time-pressure (e.g. emergency situations). “Routine tasks” describe previously planned standard procedures (e.g. standard anesthesia induction, planned surgery).

Coding of methodological factors

*Patient realism* was either coded as “Real patients” or “Simulated patients”. “Simulated patient” included a simulated or standardized patient whereas “Real patient” included real patients in clinical settings.

*Clinical performance measures* were coded either as “Outcome performance” or “Process performance”.[34,47] “Outcome performance” includes an outcome that is measured after the treatment process (e.g. infection rate, mortality). We focused only on patient-related outcomes and not on team outcomes (e.g. team satisfaction). “Process performance” describes the evaluation of the treatment process and describes how well the process was executed (e.g. adherence to guidelines through expert rating). Process performance measures are often based on official guidelines and extensive expert knowledge.[48] Thus, we assumed that process performance closely relates to patient outcomes.

*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, Schumacher, Legeret, et al.[64] | 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<br>Hands-on time within first 180 seconds |
| Brogaard, Kierkegaard, Hvidmand, et al.[65] | 2018 | To investigate the relationship between non-technical skills and clinical performance in obstetric teams | Video observation of obstetric teams (Obstetricians, obstetric 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, Kolbe, Wacker, et al.[66] | 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, Manser, Kolbe, et al.[67] | 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, Wacker, Grote, et al.[68] | 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, Min, Bridges[10] | 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, Giddings, Wilkinson, et al.[69] | 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, Mishra, Handa, et al.[70] | 2008 | To analyse the effects of surgical, aesthetic, 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 |
| Cooper, Wakelam[71] | 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 minutes | Survey about leadership behaviour using the Leadership Behaviour Description Questionnaire | Checklist based expert rating |
| Davenport, Henderson, Mosca, et al.[72] | 2007 | To measure the impact of organizational 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, Wiegmann, Henrickson, et al.[73] | 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, Chaboyer, Fairweather[74] | 2012 | To investigate how various human factors variables, extend the expected length of an operation | Live observation of surgical teams across 10 specialties | Structured observation of numbers of communication failures | Deviation from expected length of operation |

| | | | | | |
|---|---|---|---|---|---|
| Kolbe, Burtscher, Wacker, et al.[75] | 2012 | To test the relationship between speaking up and technical team performance in anaesthesia. | Observation of 2-person (nurse, resident) ad hoc anaesthesia teams performing simulated inductions of general anaesthesia with minor nonroutine events | Structured observation of speaking up behaviour | Checklist based expert rating |
| Kuenzle, Zala-Mezo, Wacker, et al.[76] | 2009 | To investigate shared leadership patterns during anaesthesia induction and to show how they are linked to team performance | Observation of 2-person (nurse, resident) ad hoc anaesthesia teams performing simulated inductions of general anaesthesia with a nonroutine event (asystole) | Structured observation of leadership behaviour | Reaction time to nonroutine event |
| Manojilovich, Antonakos, David, et al.[77] | 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<br>Bloodstream infections<br>Pressure ulcers<br>Acute physiology and chronic health evaluation score |
| Manser, Bogdanovic, Arora, et al.[78] | 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, Müller, Marquardt, et al.[79] | 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 |
| Mazzocco, Petitti, Fong, et al.[80] | 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, Catchpole, Dale, et al.[81] | 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, Hoffmann, Heimberg, et al.[82] | 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, Bristowe, Draycott, et al.[83] | 2012 | To investigate the relationship between patient satisfaction and communication | Video observation of teams (physicians, midwives) managing obstetric emergencies in secondary and tertiary maternity units | Structured observation of closed loop communication | Timely administration of magnesium sulphate |
| Siassakos, Fox, Crofts, et al.[84] | 2011 | To determine whether team performance in a simulated emergency is related to generic teamwork skills and behaviors | 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, Sexton, Lasky, et al.[85] | 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 | Compliance with Neonatal Resuscitation Program guidelines |
| Tschan, Semmer, Gautschi, et al.[86] | 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, Semmer, Gurtner, et al.[87] | 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 |
| Westli, Johnsen, Eid, et al.[88] | 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, El Bardissi, Dearani, et al.[89] | 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, Lasky, Dannemiller, et al.[90] | 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, Phillips-Bute, Petrusa, et al.[91] | 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, Fuerch, Halamek[92] | 2016 | To investigate the effect of standardized 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<br>Time to initiate positive pressure ventilation<br>Time to chest compression |

*Table 2.* Studies, effect sizes and moderator variables included in the meta-analytic database

| Authors | Year | Study goal | Participants and setting | No. of teams | Professional composition | Team familiarity | Average team size | Task type | Patient realism | Performance measure |
|---|---|---|---|---|---|---|---|---|---|---|
| Amacher, Schumacher, Legeret, et al.[64] | 2017 | .11 | Emergency medicine | 72 | Uniprofessional | Experimental | 3 | Non-routine | Simulated | Process |
| Brogaard, Kierkegaard, Hvidmand, et al.[65] | 2018 | .43 | Obstetrics | 99 | Interprofessional | Real | 5 | Non-routine | Real | Process |
| Burtscher, Kolbe, Wacker, et al.[66] | 2011 | -.27 | Anaesthesia | 31 | Interprofessional | Experimental | 2 | Routine | Simulated | Process |
| Burtscher, Manser, Kolbe, et al.[67] | 2011 | .19 | Anaesthesia | 15 | Interprofessional | Experimental | 2 | Routine & non-routine | Simulated | Process |
| Burtscher, Wacker, Grote, et al.[68] | 2010 | .07 | Anaesthesia | 22 | Interprofessional | Real | 3 | Non-routine | Real | Process |
| Carlson, Min, Bridges[10], b | 2009 | .83 | Emergency medicine | 44 | Uniprofessional | Experimental | 2.6 | Non-routine | Simulated | Process |
| Catchpole, Giddings, Wilkinson, et al.[69] | 2007 | .45† | Surgery | 24 | Interprofessional | Real | 9 | Non-routine | Real | Process |
| | 2007 | .29† | Surgery | 18 | Interprofessional | Real | 5 | Routine | Real | Process |
| Catchpole, Mishra, Handa, et al.[70] | 2008 | .36† | Surgery | 26 | Interprofessional | Real | | Routine | Real | Process |
| | 2008 | .09† | Surgery | 22 | Interprofessional | Real | | Routine | Real | Process |
| Cooper, Wakelam[71] | 1999 | .50 | General care | 20 | Interprofessional | Real | 4 | Routine | Real | Process |
| Davenport, Henderson, Mosca, et al.[72] | 2007 | .17 | Surgery | 52 | Interprofessional | Real | | Routine | Real | Outcome |
| El Bardissi, Wiegmann, Henrickson, et al.[73] | 2008 | .67 | Surgery | 31 | Interprofessional | Real | 7 | Routine | Real | Process |
| Gillespie, Chaboyer, Fairweather[74] | 2012 | .23 | Surgery | 160 | Interprofessional | Real | 6 | Routine | Real | Process |
| Kolbe, Burtscher, Wacker, et al.[75] | 2012 | .33 | Anaesthesia | 31 | Interprofessional | Real | 2 | Non-routine | Simulated | Process |
| Kuenzle, Zala-Mezo, Wacker, et al.[76] | 2009 | .56 | Anaesthesia | 12 | Interprofessional | Real | 2 | Routine | Simulated | Process |
| Manojilovich, Antonakos, David, et al.[77] | 2009 | .11 | Intensive care | 25 | Uniprofessional | Real | 36 | Routine | Real | Outcome |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Manser, Bogdanovic, Arora, et al.[78] | 2015 | .39 | Surgery | 19 | Interprofessional | Experi-mental | 5 | Routine | Simulated | Process |
| Marsch, Müller, Marquardt, et al.[79] | 2004 | .23 | Intensive care | 16 | Interprofessional | Experi-mental | 3 | Non-routine | Simulated | Process |
| Mazzocco, Petitti, Fong, et al.[80] | 2009 | .11 | Surgery | 293 | Interprofessional | Real | 6 | Routine | Real | Outcome |
| Mishra, Catchpole, Dale, et al.[81] | 2008 | .05 | Surgery | 26 | Interprofessional | Real | 6 | Routine | Real | Process |
| Schmutz, Hoffmann, Heimberg, et al.[82] | 2015 | .12 | Emergency medicine | 68 | Interprofessional | Real | 6 | Non-routine | Simulated | Process |
| Siassakos, Bristowe, Draycott, et al.[83] | 2012 | .66 | Obstetrics | 19 | Interprofessional | Real | 6 | Non-routine | Simulated | Process |
| Siassakos, Fox, Crofts, et al.[84] | 2011 | .55 | Emergency medicine/ obstetrics | 24 | Interprofessional | Experi-mental | 6 | Non-routine | Simulated | Process |
| Thomas, Sexton, Lasky, et al.[85] | 2006 | .23 | Neonatal care | 132 | Interprofessional | Real | 5 | Non-routine | Real | Process |
| Tschan, Semmer, Gautschi, et al.[86] | 2006 | .23 | Emergency medicine | 21 | Interprofessional | Experi-mental | 5 | Non-routine | Simulated | Process |
| Tschan, Semmer, Gurtner, et al.[87] | 2009 | .37 | Emergency medicine | 20 | Uniprofessional | Experi-mental | 2.65 | Non-routine | Simulated | Outcome |
| Westli, Johnsen, Eid, et al.[88] | 2010 | .18 | Emergency medicine | 27 | Interprofessional | Real | 5.1 | Non-routine | Simulated | Process |
| Wiegmann, El Bardissi, Dearani, et al.[89] | 2007 | .56 | Surgery | 31 | Interprofessional | Real | | Routine | Real | Process |
| Williams, Lasky, Dannemiller, et al.[90] | 2010 | .18 | Neonatal care | 12 | Interprofessional | Real | 5 | Non-routine | Real | Process |
| Wright, Phillips-Bute, Petrusa, et al.[91] | 2009 | .81 | General care | 9 | Uniprofessional | Experi-mental | 4 | Non-routine | Simulated | Process |
| Yamada, Fuerch, Halamek[92] | 2016 | .11 | Emergency medicine | 13 | Interprofessional | Experi-mental | 3 | Non-routine | Simulated | Process |

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

[b] Carlson, Min & Bridges has been identified as an outlier and therefore excluded from the analysis

**Statistical Analysis**

Different types of effect sizes (e.g. Odds ratio, *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.[49] and Walker.[50] 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,[44,51] we used Fisher's z score to average the multiple correlations from the same sample. 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 Alpha) or were not available at all for the majority of studies. Therefore, we report uncorrected, sample-size weighted mean correlation, its 95% confidence interval (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.[52] 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"), 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.[52] These models were calculated by the restricted maximum-likelihood estimator, an efficient and unbiased estimator.[53] 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.

The estimation of meta-analytical models including the outlier analyses were performed with the package “metafor” from the programming language and statistical environment R.[52]

## RESULTS

The online search resulted in 2002 articles (Figure 1). Based on title and abstract 67 articles were 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[69,70]). Two additional studies were identified via contacting authors directly and have been presented at conferences in the past.[65,78] This led to a total of 32 studies coming from 30 articles. Following suggestions by Viechtbauer and Cheung,[54] outliers were examined using the externally standardized residual score. One case (Carlson et al.,[10] $r = .89$, $n = 44$, standardized residual score = 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 (e.g. 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 (e.g. time until intervention). Only four studies assessed outcome performance measures (e.g. morbidity, mortality). 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 shows the relationship between teamwork and team performance. The sample-sized weighted mean correlation was .28 (95% CI: .20 – .35, $z = 6.55$, $p < .001$), indicating that teamwork is positively related to clinical performance. Results further indicated heterogeneous effect size distributions across the included samples ($Q = 53.73$, $p < .05$, $I^2 = 45.96$), 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 familiarity, team size, task type, patient realism* and *performance measures.*

The omnibus test of moderators was not significant ($F = 0.18$, $df_1 = 6$, $df_2 = 18$, $p > .20$), 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.

*Table 3.* Meta-Analytic relationships between teamwork and clinical performance

| | N | k | r | 95% CI | 80% CR | Q | $I^2$ |
|---|---|---|---|---|---|---|---|
| Overall relationship | 1,390 | 31 | .28 | [.20 ; .35] | [.09 ; .45] | 53.7 | 46.0 |
| *Team characteristics* | | | | | | | |
| Professional composition | | | | | | | |
| Interprofessional | 1,264 | 27 | .28 | [.20 ; .36] | [.09 ; .46] | 47.1* | 48.2 |
| Uniprofessional | 126 | 4 | .28 | [-.01 ; .52] | [-.04 ; .54] | 6.5 | 47.1 |
| Team familiarity | | | | | | | |
| Experimental team | 240 | 10 | .25 | [.05 ; .43] | [-.05 ; .51] | 17.2* | 47.2 |
| Real team | 1,150 | 21 | .29 | [.20 ; .37] | [.12 ; .45] | 36.2* | 45.7 |
| Team size[a] | | | | | | | |
| *Task characteristics* | | | | | | | |
| Task type | | | | | | | |
| Routine task | 766 | 14 | .27 | [.12 ; .40] | [-.01 ; .50] | 30.9* | 65.0 |
| Non-routine task | 609 | 16 | .29 | [.20 ; .39] | [.16 ; .42] | 20.5 | 24.6 |
| *Methodological factors* | | | | | | | |
| Patient realism | | | | | | | |
| Real patient | 993 | 16 | .28 | [.18 ; .38] | [.10 ; .45] | 28.7* | 49.3 |
| Simulated patient | 397 | 15 | .28 | [.13 ; .41] | [.02 ; .50] | 25.0* | 44.6 |
| Performance measures | | | | | | | |
| Outcome performance | 390 | 4 | .13 | [.03 ; .23] | [.06 ; .19] | 1.3 | 0.0 |
| Process performance | 1,000 | 27 | .30 | [.21 ; .39] | [.10 ; .49] | 45.6* | 45.6 |

*Note*. *k* = number of studies; *N* = cumulative sample size (number of teams); *r* = sample-size weighted correlation; CI = confidence interval; CR = credibility interval; *Q* = test statistic for residual heterogeneity of the models; $I^2$ = % of total variability in the effect size estimates due to heterogeneity among true effects (vs. sampling error)

[a] Team size was entered as a continuous variable, therefore, no subgroup analyses exist

## DISCUSSION

With this study, we quantified the relationship between teamwork and performance in healthcare teams. By including various contextual factors, we investigated potential contingencies that these factors might have on this relationship. The analysis of 1390 teams from 31 different studies showed that teamwork has a medium sized effect (*r*=.28;[55]) on clinical performance across various care settings. Our study is the first to investigate this relationship quantitatively with a meta-analytical procedure. We provide strong evidence that teamwork contributes considerably towards quality of care—or in other words, poor teamwork significantly increases the risk for unsafe care and even patient harm. This finding aligns with and advances previous work that explored this relationship in a qualitative way.[9,16,18,39-43]

The analysis of moderators illustrates that teamwork is related with performance under a variety of conditions. Regardless of team characteristics (professional composition, familiarity, team size) or task type (routine vs. non-routine task), teamwork influences clinical performance. Clinicians and educators from all fields should strive to maintain or increase effective teamwork. 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.

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 generalize to real life settings. 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 (e.g. 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.[19,20] Similar to our results, Hughes et al. highlighted the effectiveness of team trainings under a variety of conditions—irrespective of team composition,[19] 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.[56] A meta-analysis including teams from multiple industries (e.g. 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.[57] Most studies included in our analysis focused on rather short and intense patient care episodes (e.g. 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 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. Future research should build on recent theoretical and applied work[23,24,26,58] 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[21,23,24] 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 to 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.

Moreover, we cannot rule out a file-drawer effect, meaning that we probably could not find and include all unpublished studies, a common downside of meta-analysis.[37] Unpublished studies more often report nonsignificant results.

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. Further, 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.[59,60] Unfortunately we were not able to extract this information from the primary studies. Therefore, 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[61,62] as well as team trainings[63] 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.

**Funding statement:** 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:** All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work; no other relationships or activities that could appear to have influenced the submitted work.

**Author contributions:** 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.

**Ethics approval:** Not required

**Acknowledgements:** The authors thank Manuel Stühlinger for his help with study selection and data extraction and Walter J. Eppich, MD, MEd for a critical review and proof reading the manuscript.

**Data sharing:** Dataset is available from the corresponding author.

## REFERENCE


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. Assessing and improving teamwork in cardiac surgery. *Qual Saf Health Care* 2010;19:e29–9.

4 Williams AL, Lasky RE, Dannemiller JL, et al. Teamwork behaviours and errors during neonatal resuscitation. *BMJ Qual Saf* 2010;19:60–4.

5 Guerlain S, Turrentine FE, Bauer DT, et al. Crew resource management training for surgeons: Feasibility and impact. *Cogn Technol 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 Tech Work* 2007;139:131–274.

8 Burke CS, Salas E, Wilson-Donnelly K, et al. How to turn a team of experts into an expert medical team: guidance from the aviation and military communities. *Qual Saf Health Care* 2004;13:i96–104.

9 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.

10 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.

11 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.

12 Manojlovich M, DeCicco B. Healthy Work Environments, Nurse-Physician Communication, and Patients Outcomes. *Am J Crit Care* 2007;16:536–43.

13 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.

14 Tschan F, Semmer NK, Nägele C, et al. Task adaptive behavior and performance in groups. *Group Processes Intergroup Relat* 2000;3:367–86.

15 Manser T. Teamwork and patient safety in dynamic domains of healthcare: a review of the literature. *Acta Anaesthesiol Scand* 2009;53:143–51.

16 Fernandez Castelao 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.

17 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.

18 Flowerdew L, Brown R, Vincent C. Identifying nontechnical skills associated with safety in the emergency department: a scoping review of the literature. *Ann Emerg Med* 2012;59:386–94.

19 Hughes AM, Gregory ME, Joseph DL, et al. Saving Lives: A Meta-Analysis of

Team Training in Healthcare. *J Appl Psychol* 2016:1266-04.

20 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.

21 Fernandez R, Kozlowski SW, Shapiro MJ, et al. Toward a definition of teamwork in emergency medicine. *Acad Emerg Med* 2008;15:1104–12.

22 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, NY: Lawrence Erlbaum Associates 1997:85–108.

23 Marks MA, Mathieu JE, Zaccaro SJ. A temporally based framework and taxonomy of team processes. *Acad Manag Rev* 2001;26:356–76.

24 Reader TW, Flin R, Mearns K. Developing a team performance framework for the intensive care unit. *Crit Care Med* 2009;37:1787–93.

25 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.

26 Burke CS, Stagl KC, Salas E, et al. Understanding team adaptation: a conceptual analysis and model. *J Appl Psychol* 2006;91:1189–207.

27 Maloney MM, Bresman H, Zellmer-Bruhn ME. Contextualization and Context Theorizing in Teams Research: A Look Back and a Path Forward. *Acad Manag Ann* 2016;10:891–942.

28 Kozlowski SW, Klein KJ. A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In Klein KJ,

Kozlowski SWJ, eds. Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions. San Francisco, CA: Jossey-Bass. 2000:3-90.

29 Schofield RF, Amodeo M. Interdisciplinary teams in health care and human services settings: are they effective? *Health Soc W* 1999;24:210–9.

30 Hall P. Interprofessional teamwork: Professional cultures as barriers. *J Interprof Care* 2005;19:188–96.

31 Lillrank P. The quality of standard, routine and nonroutine processes. *Org Studies* 2003;24:215–33.

32 Eppich WJ, Oconnor LO, 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.

33 Anderson N, Ones DS, Sinangil HK, et al. Handbook of industrial, work and organizational psychology. London: Sage Publications Ltd; 2001.

34 Sonnentag S, Frese M. Performance concepts and performance theory. *Psych Manag Ind Perf* 2002;23:3–25.

35 Cooper S, Cade J. Predicting survival, in-hospital cardiac arrests: resuscitation survival variables and training effectiveness. *Resuscitation* 1997;35:17–22.

36 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.

37 Hunter JE, Schmidt FL. Methods of meta-analysis: Correcting error and bias in research findings. Thousand Oaks, CA: Sage Publications; 2004.

38 Card N. Applied Meta-Analysis for Social Science Research. New York: Guilford Publications; 2015.

39 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.

40 Ghaferi AA, Dimick JB. Importance of teamwork, communication and culture on failure-to-rescue in the elderly. *B J Surgery* 2016;103.

41 Mitchell L, Flin R. Non-technical skills of the operating theatre scrub nurse: literature review. *J Adv Nurs* 2008;63:15–24.

42 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.

43 Youngson GG. Nontechnical skills in pediatric surgery: Factors influencing operative performance. *J Ped Surgery* 2016;51:226–30.

44 Hunter JE, Schmidt FL. Methods of meta-analysis: Correcting error and bias in research findings. Thousand Oaks, CA: Sage Publications; 2004.

45 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.

46 Hunter JE, Schmidt FL. Dichotomization of continuous variables: The implications for meta-analysis. *J Appl Psychol* 1990;75:334.

47 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.

48 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.

49 Borenstein M, Hedges LV, Higgins JPT, et al. Introduction to Meta-Analysis. Hoboken, NJ: John Wiley & Sons, Ltd; 2009.

50 Walker DA. JMASM9: converting Kendall's tau for correlational or meta-analytic analyses. *J Mod Appl Stat Meth* 2003;2:525–30.

51 Lipsey MW, Wilson D. Practical meta-analysis. New York: Sage Publications, Inc; 2001.

52 Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw* 2010;36:1–48.

53 Viechtbauer W. Bias and efficiency of meta-analytic variance estimators in the random-effects model. *J Edu Behav Stat* 2005;30:261–93.

54 Viechtbauer W, Cheung MWL. Outlier and influence diagnostics for meta-analysis. *Res Synth Methods* 2010;1:112–25.

55 Cohen J. Statistical power analysis for the behavioral sciences. Hilsdale, NJ: Lawrence Erlbaum; 1988.

56 Wageman R, Baker G. Incentives and cooperation: The joint effects of task and reward interdependence on group performance. *J Organiz Behav* 1997:139–58.

57 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.

58 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. *Europ J Work Org* 2015;24:1–26.

59 Edmondson AC. Teaming: How organizations learn, innovate, and compete in the knowledge economy. San Francisco, CA: Jossey-Bass; 2012.

60 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.

61 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.

62 Schmutz JB, Eppich WJ. Promoting learning and patient care through shared reflection: A conceptual framework for team reflexivity in healthcare. *Acad Med* 2017;92:1555–63.

63 Salas E, Diaz Granados D, Weaver SJ, et al. Does Team Training Work? Principles for Health Care. *Acad Emerg Med* 2008;15:1002–9.

64 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.

65 Brogaard L, Kierkegaard O, Hvidmand L, et al. Is non-technical performance the key to high clinical performance in obstetric teams? Paper presented at the *Annual conference of the British Maternal Fetal Medicine Society*, 2018.

66 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.

67 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.

68 Burtscher MJ, Wacker J, Grote G, et al. Managing nonroutine events in anesthesia: the role of adaptive coordination. *Hum Factors* 2010;52:282–94.

69 Catchpole KR, Giddings AE, Wilkinson M, et al. Improving patient safety by identifying latent failures in successful operations. *Surgery* 2007;142:102–10.

70 Catchpole K, Mishra A, Handa A, et al. Teamwork and Error in the Operating Room. *Ann Surg* 2008;247:699–706.

71 Cooper S, Wakelam A. Leadership of resuscitation teams: "Lighthouse Leadership." *Resuscitation* 1999;42:27–45.

72 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.

73 El Bardissi 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.

74 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.

75 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.

76 Künzle B, Zala-Mezo E, Wacker J. Leadership in anaesthesia teams: the most effective leadership is shared. *Qual Saf Health Care* 2010;19:1–6.

77 Manojlovich M, DeCicco B. Healthy Work Environments, Nurse-Physician Communication, and Patients Outcomes. *Am J Crit Care* 2007;16:536–43.

78 Manser T, Bogdanovic J, Arora S, et al. Effective Team Management as a Critical Coordination Skill of Surgical Leaders. Paper presented at the *9th International Meeting for Behavioural Science in Surgery & Acute Care Settings*, 2015.

79 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.

80 Mazzocco K, Petitti DB, Fong KT, et al. Surgical team behaviors and patient outcomes. *Am J Surg* 2009;197:678–85.

81 Mishra A, Catchpole K, Dale T, et al. The influence of non-technical performance on technical outcome in laparoscopic cholecystectomy. *Surg Endosc* 2007;22:68–73.

82 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.

83 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;32:41.

84 Siassakos D, Fox R, Crofts JF, et al. The management of a simulated emergency: Better teamwork, better performance. *Resuscitation*

85 Thomas EJ, Sexton JB, Lasky RE, et al. Teamwork and quality during neonatal care in the delivery room. *J Perinatol* 2006;26:163–9.

86 Tschan F, Semmer NK, Gautschi, et al. Leading to Recovery: Group Performance and Coordinative Activities in Medical Emergency Driven Groups. *Hum Perf* 2006;19:277–304.

87 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.

88 Westli H, Johnsen B, 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.

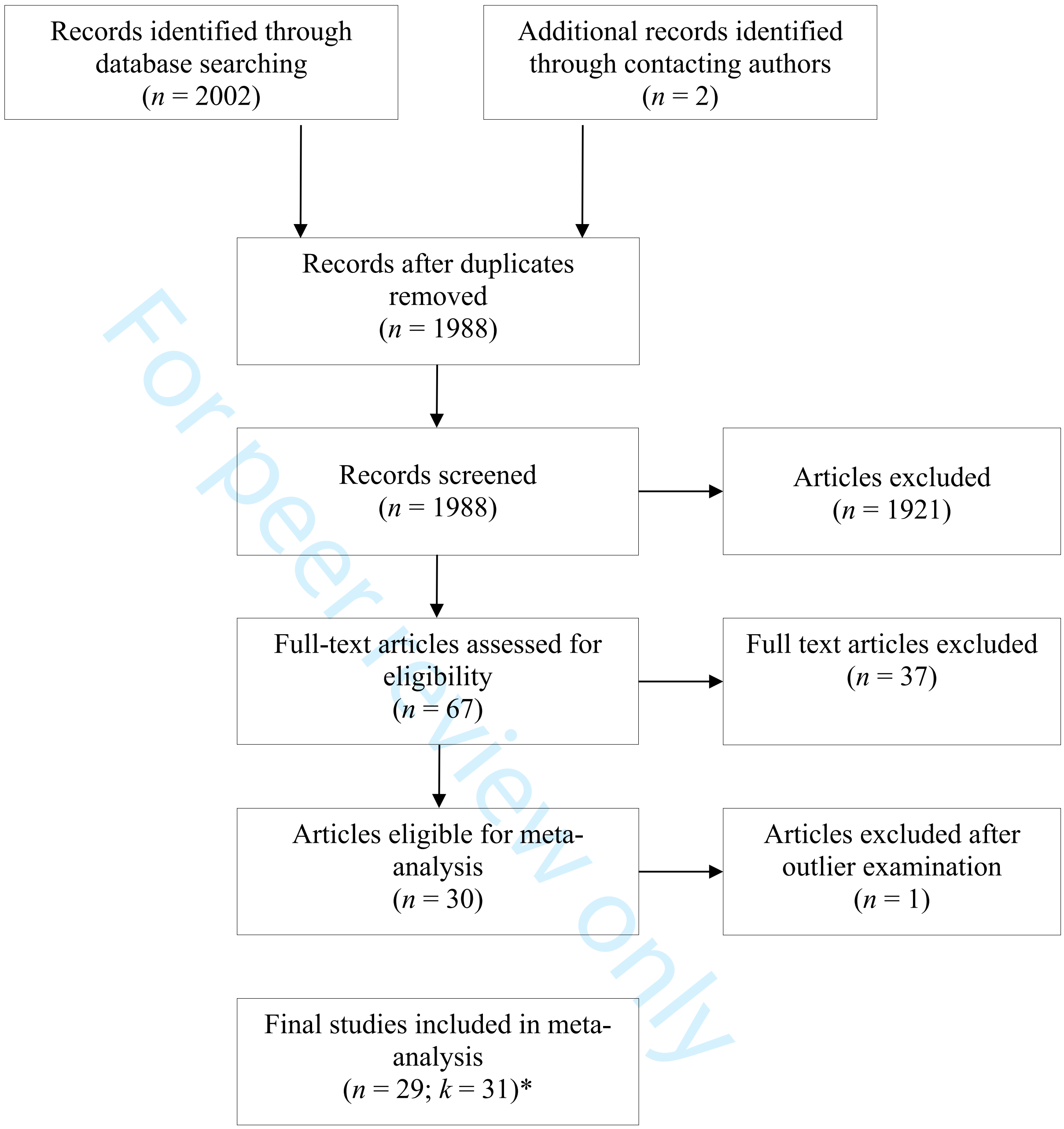
89 Wiegmann DA, ElBardissi AW, Dearani JA. Disruptions in surgical flow and their relationship to surgical errors: an exploratory investigation. *Surgery* 2007;142:658–65.

90 Williams AL, Lasky RE, Dannemiller JL, et al. Teamwork behaviours and errors during neonatal resuscitation. *Qual Saf Health Care* 2010;19:60–4.

91 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.

92 Yamada NK, Fuerch JH, Halamek LP. Impact of standardized communication techniques on errors during simulated neonatal resuscitation. *AM J Perinatol* 2016;33:385–92.

Figure 1. Systematic literature search.




*The search identified 31 studies published in 29 articles

## SUPPLEMENTARY FILE

Article: **How effective is teamwork really? The relationship between teamwork and performance in healthcare teams: A systematic review and meta-analysis**

Jan B. Schmutz, PhD, Laurenz L. Meier, PhD, Tanja Manser, PhD

## Search terms used for PubMed search

(Teamwork[All Fields]) OR (team[All Fields] AND coordination[All Fields]) OR (team[All Fields] AND "decision making"[All Fields]) OR ((team[All Fields]) AND (communication[MeSH Terms]) OR (communication[All Fields])) OR (team[All Fields] AND leadership[All Fields]) AND (patient safety[MeSH Terms])

# PRISMA 2009 Checklist

| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| **TITLE** | | | |
| Title | 1 | Identify the report as a systematic review, meta-analysis, or both. | 1 |
| **ABSTRACT** | | | |
| Structured summary | 2 | Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number. | 2 |
| **INTRODUCTION** | | | |
| Rationale | 3 | Describe the rationale for the review in the context of what is already known. | 4-8 |
| Objectives | 4 | Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS). | 4-5 |
| **METHODS** | | | |
| Protocol and registration | 5 | Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number. | n/a |
| Eligibility criteria | 6 | Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. | 9-10 |
| Information sources | 7 | Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched. | 9 |
| Search | 8 | Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated. | 9 and supplemental material |
| Study selection | 9 | State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). | 9-10 |
| Data collection process | 10 | Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators. | 10 |
| Data items | 11 | List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made. | 9-12 |
| Risk of bias in individual studies | 12 | Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis. | 20 |
| Summary measures | 13 | State the principal summary measures (e.g., risk ratio, difference in means). | 13-19 / 23 |

# PRISMA 2009 Checklist

| | | | |
|---|---|---|---|
| Synthesis of results | 14 | Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $I^2$) for each meta-analysis. | 20-23 |



| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| Risk of bias across studies | 15 | Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies). | 20,26 |
| Additional analyses | 16 | Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified. | 22-23 |
| **RESULTS** | | | |
| Study selection | 17 | Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram. | 21 |
| Study characteristics | 18 | For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations. | 18-19 |
| Risk of bias within studies | 19 | Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12). | 18-19, 20 |
| Results of individual studies | 20 | For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot. | 13-19 |
| Synthesis of results | 21 | Present results of each meta-analysis done, including confidence intervals and measures of consistency. | 23 |
| Risk of bias across studies | 22 | Present results of any assessment of risk of bias across studies (see Item 15). | 23 |
| Additional analysis | 23 | Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]). | N/A |
| **DISCUSSION** | | | |
| Summary of evidence | 24 | Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers). | 24-27 |
| Limitations | 25 | Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias). | 26-27 |
| Conclusions | 26 | Provide a general interpretation of the results in the context of other evidence, and implications for future research. | 27 |
| **FUNDING** | | | |
| Funding | 27 | Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. | 28 |

# PRISMA 2009 Checklist

*From:* Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

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



| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2018-028280.R1 |
| Article Type: | Research |
| Date Submitted by the Author: | 01-Apr-2019 |





| | |
|---|---|
| Complete List of Authors: | Schmutz, Jan B.; Northwestern University, Department of Communication Studies<br>Meier, Laurenz ; University of Neuchâtel , Occupational Psychology and Organizations<br>Manser, Tanja; University of Applied Sciences and Arts Northwestern Switzerland |





| | |
|---|---|
| <b>Primary Subject Heading</b>: | Communication |
| Secondary Subject Heading: | Communication |
| Keywords: | teamwork, non-technical skills, communication, meta-analysis, teams |

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

**Jan B. Schmutz, PhD**

Researcher at the Department of Communication Studies, Northwestern University, Evanston, USA

Postdoctoral Researcher and Lecturer at the Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland

jan.schmutz@northwestern.edu

**Laurenz L. Meier, PhD**

Professor at the Institute of Work and Organizational Psychology, University of Neuchâtel, Neuchâtel, Switzerland

laurenz.meier@unine.ch

**Tanja Manser, PhD**

Director of the FHNW School of Applied Psychology, University of Applied Sciences and Arts

Northwestern Switzerland, Olten, Switzerland

tanja.manser@fhnw.ch

Word count: 4647

Keywords: teamwork, non-technical skills, communication, teams, meta-analysis



**Correspondence to**

Jan B. Schmutz
ATLAS Lab
Department of Communication Studies
Northwestern University
Frances Searle Building 1.220
2240 Campus Drive
Evanston, IL 60208
United States
Phone 847-491-7023
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.

**Methods** Data sources were searched up to June 2018 and included PubMed, manual backward search of relevant reviews, manual backward and forward search of studies included in the meta-analysis and contacting of selected authors via e-mail. Studies were included if they reported a relationship between a teamwork process and a performance measure. Moderator variables (i.e. 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 PRISMA 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 $r$ = .28, 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 organizations should recognize the value of teamwork and emphasize approaches that maintain and improve teamwork for the benefit of their patients.

## ARTICLE SUMMARY

### Strengths and limitations of this study

- This systematic review evaluates available studies investigating the effectiveness of teamwork processes.
- 31 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.

## 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 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 systems[1-4] 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 (e.g. 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 organization 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 (e.g. 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 organizations.

In the following we will first establish an operational definition of teamwork, elaborate on relevant contextual factors, and present our respective meta-analytic 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, *e)* distributed expertise and *f)* clearly assigned roles and responsibilities.[20,21] Based on this definition it becomes clear that teams must dynamically share information and resources amongst 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 (e.g. 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 (i.e. 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* (i.e. 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 emphasizes 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 criticized because of its simplicity.[30] However, it is still the most popular

framework to date and helps to systematize 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 hypothesize 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 (i.e. professional composition, team familiarity, team size); (b) task type (i.e. routine vs. non-routine tasks); (c) two methodological factors related to patient realism (i.e. simulated vs. real) and the type of performance measures used (i.e. 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 to 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 to smaller teams.

*Task type.* Routine situations are characterized by repetitive and unvarying actions (e.g. 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 standardized 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 generalize 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- and outcome-related aspects of performance.[37,38] Process performance measures are action-related aspects and refer to adequate behaviour during procedures (e.g. adhering to guidelines), making them easier to assess. Outcome performance measures (e.g. 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 (e.g. 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 (i.e. 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 (i.e. high process performance) the patient can still die (i.e. 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 PRISMA statement[40] 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 data base PubMed (no limit was placed on date of publication, last search June 2018) using the key words *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 8 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 e-mail 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 (e.g. coordination, communication). In addition, studies needed to investigate the relationship between at least one teamwork process and a performance measure (e.g. patient outcome). When studies reported multiple estimates of the same relationship from

the same sample (e.g. 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 dynamics of teamwork over a longer period of time are different. All articles included in this meta-analysis are listed in Table 1 and Table 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-analytic integrations.[48,49]

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. 20% of the studies were rated by both coders. Intercoder 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 (e.g. 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 behavior often under time-pressure (e.g. emergency situations). "Routine tasks" describe previously planned standard procedures (e.g. standard anesthesia 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".[38,50] "Outcome performance" includes an outcome that is measured after the treatment process (e.g. infection rate, mortality). We focused only on patient-related outcomes and not on team outcomes (e.g. team satisfaction). "Process performance" describes the evaluation of the treatment process and describes how well the process was executed (e.g.

adherence to guidelines through expert rating). Process performance measures are often based on official guidelines and extensive expert knowledge.[51] Thus, we assumed that process performance closely relates to patient outcomes.

**Statistical Analysis**

Different types of effect sizes (e.g. Odds ratio, *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.[52] and Walker.[53] 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,[41,54] we used Fisher's z score to average the multiple correlations from the same sample. 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 Alpha) or were not available at all for the majority of studies. Therefore, we report uncorrected, sample-size weighted mean correlation, its 95% confidence interval (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.[55] 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"), 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.[55] These models were calculated by the restricted maximum-likelihood estimator, an efficient and unbiased estimator.[56] 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.[57] The results indicate that there is no asymmetry in the funnel plot (z = 1.79, p = .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.[55]

**Patient and public involvement**

Patients and public were not involved in this study.

## RESULTS

The online search resulted in 2002 articles (Figure 1). Based on title and abstract 67 articles were 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[58,59]). Two additional studies were identified via contacting authors directly and have been presented at conferences in the past.[60,61] This led to a total of 32 studies coming from 30 articles. One case (Carlson et al.,[9] $r = .89$, $n = 44$, standardized residual score = 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 (e.g. 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 (e.g. time until intervention). Only four studies assessed outcome performance measures. Measures included accuracy of diagnosis, postoperative 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 .28 (95% CI: .20 to .35, $z = 6.55$, $p < .001$), indicating that teamwork is positively related to clinical performance. Results further indicated heterogeneous effect size distributions across the included samples ($Q = 53.73$, $p < .05$, $I^2 =$

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 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, $p$ > .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*=.28; [62,63]) 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*=.28 might not seem very high. However, we would like to highlight that *r*=.28 is considered a medium sized effect[62,63] and should not be underestimated. To better illustrate what this effect means we transformed the correlation into an odds ratio (OR) of 2.8.[52] 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 (e.g. mortality, morbidity) or are closely related to patient outcomes (e.g. 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.[64] 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 rigor. Guidelines on how to develop performance assessment tools for specific clinical scenarios exist and need to be accounted for.[51,65,66] 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 (e.g. emergencies).[67] Our results suggest that team trainings should not only focus on non-routine situations like emergencies but also on routine situations (e.g. 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 generalize to real life settings. 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 (e.g. 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.[68] A meta-analysis including teams from multiple industries (e.g. 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.[69] Most studies included in our analysis focused on rather short and intense patient care episodes (e.g. 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 data base PubMed.

Future research should build on recent theoretical and applied work[24,26,28,70] 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 to 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.[71,72]

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.[73] 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[74,75] as well as team trainings[76] 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.

**Funding statement:** 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:** All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work; no other relationships or activities that could appear to have influenced the submitted work.

**Author contributions:** 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.

**Ethics approval:** Not required

**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 proof reading the manuscript.

**Data sharing:** Meta-analysis dataset is available from the corresponding author.

**REFERENCE**

1 Thomas EJ. Improving teamwork in healthcare: current approaches and the path forward. BMJ Quality & Safety 2011;20:647–50.

2 Salas E, Rosen MA. Building high reliability teams: progress and some reflections on teamwork training. BMJ Quality & Safety 2013;22:369–73.

3 Schraagen JM, Schouten T, Smit M, Haas F, van der Beek D, van de Ven J, et al. A prospective study of paediatric cardiac surgical microsystems: assessing the relationships between non-routine events, teamwork and patient outcomes. BMJ Quality & Safety 2011;20:599–603.

4 Williams AL, Lasky RE, Dannemiller JL, Andrei AM, Thomas EJ. Teamwork behaviours and errors during neonatal resuscitation. BMJ Quality & Safety 2010;19:60–4.

5 Guerlain S, Turrentine FE, Bauer DT, Calland JF, Adams R. Crew resource management training for surgeons: Feasibility and impact. Cognition, Technology & Work 2008;10:255–64.

6 McCulloch P, Mishra A, Handa A, Dale T, Hirst G, Catchpole K. 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, Youngson G, Mitchell A, Rowley D, et al. Debriefing surgeons on non-technical skills (NOTSS). Cogn Tech 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, Manser T. Effective coordination in medical emergency teams: The moderating role of task type. European Journal of Work and Organizational Psychology 2015;24:761–76.

13 Tschan F, Semmer NK, Nägele C, Gurtner A. 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 Castelao E, Russo SG, Riethmüller M, Boos M. 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, Mendez-Tellez PA, Dwyer C, Wyskiel R, et al. A systematic review of behavioural marker systems in healthcare: what do we know about their attributes, validity and application? BMJ Quality & Safety 2014;23:1031–9.

17 Flowerdew L, Brown R, Vincent C, Woloshynowych M. 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, Sonesh SC, Marlow SL, Lacerenza CN, et al. Saving Lives: A Meta-Analysis of Team Training in Healthcare. J Appl Psychol 2016:1266–304.

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. Theoretical Issues in Ergonomics Science 2007;8:381–94.

21 Salas E, Stagl KC, Burke CS, Goodwin GF. Fostering team effectiveness in organizations: Toward an integrative theoretical framework. vol. 52, 2007, p. 185.

22 Fernandez R, Kozlowski SW, Shapiro MJ, Salas E. 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, editors. Team Performance Assessment and Measurement: Theory, methods, and applications, New York: Team performance assessment and measurement: Theory, methods, and applications; 1997, pp. 85–108.

24 Marks MA, Mathieu JE, Zaccaro SJ. A temporally based framework and taxonomy of team processes. The Academy of Management Review 2001;26:356–76.

25 Dietz AS, Pronovost PJ, Mendez-Tellez PA, Wyskiel R, Marsteller JA, Thompson DA, 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, Cuthbertson BH. Developing a team performance framework for the intensive care unit. Crit Care Med 2009;37:1787–93.

27 Ilgen DR, Hollenbeck JR, Johnson M, Jundt D. 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, Pierce L, Kendall D. Understanding team adaptation: a conceptual analysis and model. J Appl Psychol 2006;91:1189–207.

29 Lemieux-Charles L. 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, Beaver GR. Contextualization and Context Theorizing in Teams Research: A Look Back and a Path Forward. The 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 & Social Work 1999;24:210–9.

34 Hall P. Interprofessional teamwork: Professional cultures as barriers. Journal of Interprofessional Care 2005;19:188–96.

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, editors. Essential Simulation in Clinical Education, Oxford, UK: John Wiley & Sons, Ltd; 2013, pp. 213–34.

37 Anderson N, Ones DS, Sinangil HK, Viswesvaran C. Handbook of industrial, work and organizational psychology: Personnel psychology. vol. 1. London: Sage Publications Ltd; 2001.

38 Sonnentag S, Frese M. Performance concepts and performance theory. 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, Mulrow C, Gøtzsche PC, Ioannidis JPA, 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 FL. Methods of meta-analysis: Correcting error and bias in research findings. Thousand Oaks, 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, Glavin RJ, Maran NJ. 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. Importance of teamwork, communication and culture on failure-to-rescue in the elderly. British Journal of Surgery 2016;103.

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, Alves C, Fragata I, Fragata J. 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. Journal of Pediatric Surgery 2016;51:226–30.

48 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.

49 Hunter JE, Schmidt FL. Dichotomization of continuous variables: The implications for meta-analysis. Journal of Applied Psychology 1990;75:334.

50 Campbell JP. Modeling the performance prediction problem in industrial and organizational psychology. In: Dunnette MD, Hough LM, editors. Handbook of industrial and organizational psychology, Palo Alto, CA: Consulting Psychologists Press; 1990, pp. 687–732.

51 Schmutz J, Eppich WJ, Hoffmann F, Heimberg E, Manser T. Five steps to develop checklists for evaluating clinical performance: An integrative approach. Acad Med 2014;89:996–1005.

52 Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis. Hoboken, NJ: John Wiley & Sons, Ltd; 2009.

53 Walker DA. JMASM9: converting Kendall's tau for correlational or meta-analytic analyses. J Mod Appl Stat Meth 2003;2:525–30.

54 Lipsey MW, Wilson D. Practical meta-analysis. New York: Sage Publications, Inc; 2001.

55 Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw 2010;36:1–48.

56 Viechtbauer W. Bias and efficiency of meta-analytic variance estimators in the random-effects model. Journal of Educational and Behavioral Statistics 2005;30:261–93.

57 Sterne JA, Egger M. Regression methods to detect publication and other bias in meta-analysis. Publication bias in meta-analysis: Prevention, assessment and adjustments, Chichester: John Wiley & Sons; 2006, pp. 99–110.

58 Catchpole KR, Giddings AE, Wilkinson M, Hirst G, Dale T, de Leval MR. Improving patient safety by identifying latent failures in successful operations. Surgery 2007;142:102–10.

59 Catchpole K, Mishra A, Handa A, McCulloch P. Teamwork and Error in the Operating Room. Ann Surg 2008;247:699–706.

60 Brogaard L, Kierkegaard O, Hvidmand L, Jensen KR, Uldbjerg N, Musaeus P, et al. 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., Arora S, Sevdalis N., Schmutz, J. Surgeon's team management skills predict clinical performance. Paper presented at the Annual Meeting of the Society in Europe for Simulation Applied to Medecine (SESAM), Lisbon, Portugal, 2015.

62 Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum; 1988.

63 Bosco FA, Aguinis H, Singh K, Field JG, Pierce CA. Correlational effect size benchmarks. Journal of Applied Psychology 2015;100:431.

64 Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat A-HS, Dellinger EP, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. N

Engl J Med 2009;360:491–9.

65 Gaba DM, Howard SK, Flanagan B, Smith BE, Fish KJ, Botney R. Assessment of Clinical Performance during Simulated Crises Using Both Technical and Behavioral Ratings. Anesthesiology 1998;89:8–18.

66 Rosen MA, Pronovost PJ. Advancing the use of checklists for evaluating performance in health care. Acad Med 2014;89.

67 Gaba DM, Howard SK, Fish KJ, Smith BE, Sowb YA. Simulation-based training in anesthesia crisis resource management (ACRM): a decade of experience. Simulation & Gaming 2001;32:175–93.

68 Wageman R, Baker G. Incentives and cooperation: The joint effects of task and reward interdependence on group performance. J Organiz Behav 1997:139–58.

69 LePine JA, Piccolo RF, Jackson CL, Mathieu JE, Saul JR. A meta-analysis of teamwork processes: tests of a multidimensional model and relationships with team effectiveness criteria. Pers Psychol 2008;61:273–307.

70 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. European Journal of Work and Organizational Psychology 2015;24:1–26.

71 Edmondson AC. Teaming: How organizations learn, innovate, and compete in the knowledge economy. San Francisco, CA: Jossey-Bass; 2012.

72 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.

73 Henrickson Parker, S., Schmutz, J. B., & Manser, T. Training needs for adaptive coordination: Utilizing task analysis to identify coordination requirements in three

different clinical settings. Gr & Org Manag 2018;43:504-527.

74 Eppich WJ, Mullan PC, Brett-Fleegler M, Cheng A. “Let's Talk About It”: Translating Lessons From Health Care Simulation to Clinical Event Debriefings and Coaching Conversations. Clinical Pediatric Emergency Medicine 2016;17:200–11.

75 Schmutz JB, Eppich WJ. Promoting learning and patient care through shared reflection: A conceptual framework for team reflexivity in healthcare. Acad Med 2017;92:1555–63.

76 Salas E, DiazGranados D, Weaver SJ, King H. Does Team Training Work? Principles for Health Care. Acad Emerg Med 2008;15:1002–9.

77 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.

78 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.

79 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.

80 Burtscher MJ, Wacker J, Grote G, et al. Managing nonroutine events in anesthesia: the role of adaptive coordination. *Hum Factors* 2010;52:282–94.

81 Cooper S, Wakelam A. Leadership of resuscitation teams: “Lighthouse Leadership.” *Resuscitation* 1999;42:27–45.

82 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.

83 El Bardissi 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.

84 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.

85 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.

86 Künzle B, Zala-Mezo E, Wacker J. Leadership in anaesthesia teams: the most effective leadership is shared. *Qual Saf Health Care* 2010;19:1–6.

87 Manojlovich M, DeCicco B. Healthy Work Environments, Nurse-Physician Communication, and Patients Outcomes. *Am J Crit Care* 2007;16:536–43.

88 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.

89 Mazzocco K, Petitti DB, Fong KT, et al. Surgical team behaviors and patient outcomes. *Am J Surg* 2009;197:678–85.

90 Mishra A, Catchpole K, Dale T, et al. The influence of non-technical performance on technical outcome in laparoscopic cholecystectomy. *Surg Endosc* 2007;22:68–73.

91 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.

92 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;32:41.

93 Siassakos D, Fox R, Crofts JF, et al. The management of a simulated emergency: Better teamwork, better performance. *Resuscitation*

94 Thomas EJ, Sexton JB, Lasky RE, et al. Teamwork and quality during neonatal care in the delivery room. *J Perinatol* 2006;26:163–9.

95 Tschan F, Semmer NK, Gautschi, et al. Leading to Recovery: Group Performance and Coordinative Activities in Medical Emergency Driven Groups. *Hum Perf* 2006;19:277–304.

96 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.

97 Westli H, Johnsen B, 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.

98 Wiegmann DA, ElBardissi AW, Dearani JA. Disruptions in surgical flow and their relationship to surgical errors: an exploratory investigation. *Surgery* 2007;142:658–65.

99 Williams AL, Lasky RE, Dannemiller JL, et al. Teamwork behaviours and errors during neonatal resuscitation. *Qual Saf Health Care* 2010;19:60–4.

100 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.

101 Yamada NK, Fuerch JH, Halamek LP. Impact of standardized communication techniques on errors during simulated neonatal resuscitation. *AM J Perinatol* 2016;33:385–92.

*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, Schumacher, Legeret, et al.[77] | 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<br>Hands-on time within first 180 seconds |
| Brogaard, Kierkegaard, Hvidmand, et al.[60] | 2018 | To investigate the relationship between non-technical skills and clinical performance in obstetric teams | Video observation of obstetric teams (Obstetricians, obstetric 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, Kolbe, Wacker, et al.[78] | 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, Manser, Kolbe, et al.[79] | 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, Wacker, Grote, et al.[80] | 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, Min, Bridges[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, Giddings, Wilkinson, et al.[58] | 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, Mishra, Handa, et al.[59] | 2008 | To analyse the effects of surgical, aesthetic, 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 |
| Cooper, Wakelam[81] | 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 minutes | Survey about leadership behaviour using the Leadership Behaviour Description Questionnaire | Checklist based expert rating |
| Davenport, Henderson, Mosca, et al.[82] | 2007 | To measure the impact of organizational 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, Wiegmann, Henrickson, et al.[83] | 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, Chaboyer, Fairweather[84] | 2012 | To investigate how various human factors variables, extend the expected length of an operation | Live observation of surgical teams across 10 specialties | Structured observation of numbers of communication failures | Deviation from expected length of operation |

| | | | | | |
|---|---|---|---|---|---|
| Kolbe, Burtscher, Wacker, et al.[85] | 2012 | To test the relationship between speaking up and technical team performance in anaesthesia. | Observation of 2-person (nurse, resident) ad hoc anaesthesia teams performing simulated inductions of general anaesthesia with minor nonroutine events | Structured observation of speaking up behaviour | Checklist based expert rating |
| Kuenzle, Zala-Mezo, Wacker, et al.[86] | 2009 | To investigate shared leadership patterns during anaesthesia induction and to show how they are linked to team performance | Observation of 2-person (nurse, resident) ad hoc anaesthesia teams performing simulated inductions of general anaesthesia with a nonroutine event (asystole) | Structured observation of leadership behaviour | Reaction time to nonroutine event |
| Manojilovich, Antonakos, David, et al.[87] | 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<br>Bloodstream infections<br>Pressure ulcers<br>Acute physiology and chronic health evaluation score |
| Manser, Bogdanovic, Clack, et al. [61] | 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, Müller, Marquardt, et al.[88] | 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 |
| Mazzocco, Petitti, Fong, et al.[89] | 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, Catchpole, Dale, et al.[90] | 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, Hoffmann, Heimberg, et al.[91] | 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, Bristowe, Draycott, et al.[92] | 2012 | To investigate the relationship between patient satisfaction and communication | Video observation of teams (physicians, midwives) managing obstetric emergencies in secondary and tertiary maternity units | Structured observation of closed loop communication | Timely administration of magnesium sulphate |
| Siassakos, Fox, Crofts, et al.[93] | 2011 | To determine whether team performance in a simulated emergency is related to generic teamwork skills and behaviors | 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, Sexton, Lasky, et al.[94] | 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 | Compliance with Neonatal Resuscitation Program guidelines |
| Tschan, Semmer, Gautschi, et al.[95] | 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, Semmer, Gurtner, et al.[96] | 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 |
| Westli, Johnsen, Eid, et al.[97] | 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, El Bardissi, Dearani, et al.[98] | 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, Lasky, Dannemiller, et al.[99] | 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, Phillips-Bute, Petrusa, et al.[100] | 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, Fuerch, Halamek[101] | 2016 | To investigate the effect of standardized 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<br>Time to initiate positive pressure ventilation<br>Time to chest compression |

*Table 2.* Studies, effect sizes and moderator variables included in the meta-analytic database

| Authors | Year | Study goal | Setting | No. of teams | Professional composition | Team familiarity | Average team size | Task type | Patient realism | Performance measure |
|---|---|---|---|---|---|---|---|---|---|---|
| Amacher, Schumacher, Legeret, et al.[77] | 2017 | .11 | Emergency medicine | 72 | Uniprofessional | Experimental | 3 | Non-routine | Simulated | Process |
| Brogaard, Kierkegaard, Hvidmand, et al.[60] | 2018 | .43 | Obstetrics | 99 | Interprofessional | Real | 5 | Non-routine | Real | Process |
| Burtscher, Kolbe, Wacker, et al.[78] | 2011 | -.27 | Anaesthesia | 31 | Interprofessional | Experimental | 2 | Routine | Simulated | Process |
| Burtscher, Manser, Kolbe, et al.[79] | 2011 | .19 | Anaesthesia | 15 | Interprofessional | Experimental | 2 | Routine & non-routine | Simulated | Process |
| Burtscher, Wacker, Grote, et al.[80] | 2010 | .07 | Anaesthesia | 22 | Interprofessional | Real | 3 | Non-routine | Real | Process |
| Carlson, Min, Bridges[9][b] | 2009 | .83 | Emergency medicine | 44 | Uniprofessional | Experimental | 2.6 | Non-routine | Simulated | Process |
| Catchpole, Giddings, Wilkinson, et al.[58] | 2007 | .45† | Surgery | 24 | Interprofessional | Real | 9 | Non-routine | Real | Process |
| | 2007 | .29† | Surgery | 18 | Interprofessional | Real | 5 | Routine | Real | Process |
| Catchpole, Mishra, Handa, et al.[59] | 2008 | .36† | Surgery | 26 | Interprofessional | Real | | Routine | Real | Process |
| | 2008 | .09† | Surgery | 22 | Interprofessional | Real | | Routine | Real | Process |
| Cooper, Wakelam[81] | 1999 | .50 | General care | 20 | Interprofessional | Real | 4 | Routine | Real | Process |
| Davenport, Henderson, Mosca, et al.[82] | 2007 | .17 | Surgery | 52 | Interprofessional | Real | | Routine | Real | Outcome |
| El Bardissi, Wiegmann, Henrickson, et al.[83] | 2008 | .67 | Surgery | 31 | Interprofessional | Real | 7 | Routine | Real | Process |
| Gillespie, Chaboyer, Fairweather[84] | 2012 | .23 | Surgery | 160 | Interprofessional | Real | 6 | Routine | Real | Process |
| Kolbe, Burtscher, Wacker, et al.[85] | 2012 | .33 | Anaesthesia | 31 | Interprofessional | Real | 2 | Non-routine | Simulated | Process |
| Kuenzle, Zala-Mezo, Wacker, et al.[86] | 2009 | .56 | Anaesthesia | 12 | Interprofessional | Real | 2 | Routine | Simulated | Process |
| Manojilovich, Antonakos, David, et al.[87] | 2009 | .11 | Intensive care | 25 | Uniprofessional | Real | 36 | Routine | Real | Outcome |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Manser, Bogdanovic, Clack, et al.[61] | 2015 | .39 | Surgery | 19 | Interprofessional | Experi-mental | 5 | Routine | Simulated | Process |
| Marsch, Müller, Marquardt, et al.[88] | 2004 | .23 | Intensive care | 16 | Interprofessional | Experi-mental | 3 | Non-routine | Simulated | Process |
| Mazzocco, Petitti, Fong, et al.[89] | 2009 | .11 | Surgery | 293 | Interprofessional | Real | 6 | Routine | Real | Outcome |
| Mishra, Catchpole, Dale, et al.[90] | 2008 | .05 | Surgery | 26 | Interprofessional | Real | 6 | Routine | Real | Process |
| Schmutz, Hoffmann, Heimberg, et al.[91] | 2015 | .12 | Emergency medicine | 68 | Interprofessional | Real | 6 | Non-routine | Simulated | Process |
| Siassakos, Bristowe, Draycott, et al.[92] | 2012 | .66 | Obstetrics | 19 | Interprofessional | Real | 6 | Non-routine | Simulated | Process |
| Siassakos, Fox, Crofts, et al.[93] | 2011 | .55 | Emergency medicine/ obstetrics | 24 | Interprofessional | Experi-mental | 6 | Non-routine | Simulated | Process |
| Thomas, Sexton, Lasky, et al.[94] | 2006 | .23 | Neonatal care | 132 | Interprofessional | Real | 5 | Non-routine | Real | Process |
| Tschan, Semmer, Gautschi, et al.[95] | 2006 | .23 | Emergency medicine | 21 | Interprofessional | Experi-mental | 5 | Non-routine | Simulated | Process |
| Tschan, Semmer, Gurtner, et al.[96] | 2009 | .37 | Emergency medicine | 20 | Uniprofessional | Experi-mental | 2.65 | Non-routine | Simulated | Outcome |
| Westli, Johnsen, Eid, et al.[97] | 2010 | .18 | Emergency medicine | 27 | Interprofessional | Real | 5.1 | Non-routine | Simulated | Process |
| Wiegmann, El Bardissi, Dearani, et al.[98] | 2007 | .56 | Surgery | 31 | Interprofessional | Real | | Routine | Real | Process |
| Williams, Lasky, Dannemiller, et al.[99] | 2010 | .18 | Neonatal care | 12 | Interprofessional | Real | 5 | Non-routine | Real | Process |
| Wright, Phillips-Bute, Petrusa, et al.[100] | 2009 | .81 | General care | 9 | Uniprofessional | Experi-mental | 4 | Non-routine | Simulated | Process |
| Yamada, Fuerch, Halamek[101] | 2016 | .11 | Emergency medicine | 13 | Interprofessional | Experi-mental | 3 | Non-routine | Simulated | Process |

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

[b] Carlson, Min & Bridges has been identified as an outlier and therefore excluded from the analysis

*Table 3.* Meta-Analytic relationships between teamwork and clinical performance

| | N | k | r | 95% CI | 80% CR | Q | $I^2$ |
|---|---|---|---|---|---|---|---|
| Overall relationship | 1,390 | 31 | .28 | [.20 ; .35] | [.09 ; .45] | 53.7 | 46.0 |
| *Team characteristics* | | | | | | | |
| Professional composition | | | | | | | |
| Interprofessional | 1,264 | 27 | .28 | [.20 ; .36] | [.09 ; .46] | 47.1* | 48.2 |
| Uniprofessional | 126 | 4 | .28 | [-.01 ; .52] | [-.04 ; .54] | 6.5 | 47.1 |
| Team familiarity | | | | | | | |
| Experimental team | 240 | 10 | .25 | [.05 ; .43] | [-.05 ; .51] | 17.2* | 47.2 |
| Real team | 1,150 | 21 | .29 | [.20 ; .37] | [.12 ; .45] | 36.2* | 45.7 |
| Team size[a] | | | | | | | |
| *Task characteristics* | | | | | | | |
| Task type | | | | | | | |
| Routine task | 766 | 14 | .27 | [.12 ; .40] | [-.01 ; .50] | 30.9* | 65.0 |
| Non-routine task | 609 | 16 | .29 | [.20 ; .39] | [.16 ; .42] | 20.5 | 24.6 |
| *Methodological factors* | | | | | | | |
| Patient realism | | | | | | | |
| Real patient | 993 | 16 | .28 | [.18 ; .38] | [.10 ; .45] | 28.7* | 49.3 |
| Simulated patient | 397 | 15 | .28 | [.13 ; .41] | [.02 ; .50] | 25.0* | 44.6 |
| Performance measures | | | | | | | |
| Outcome performance | 390 | 4 | .13 | [.03 ; .23] | [.06 ; .19] | 1.3 | 0.0 |
| Process performance | 1,000 | 27 | .30 | [.21 ; .39] | [.10 ; .49] | 45.6* | 45.6 |

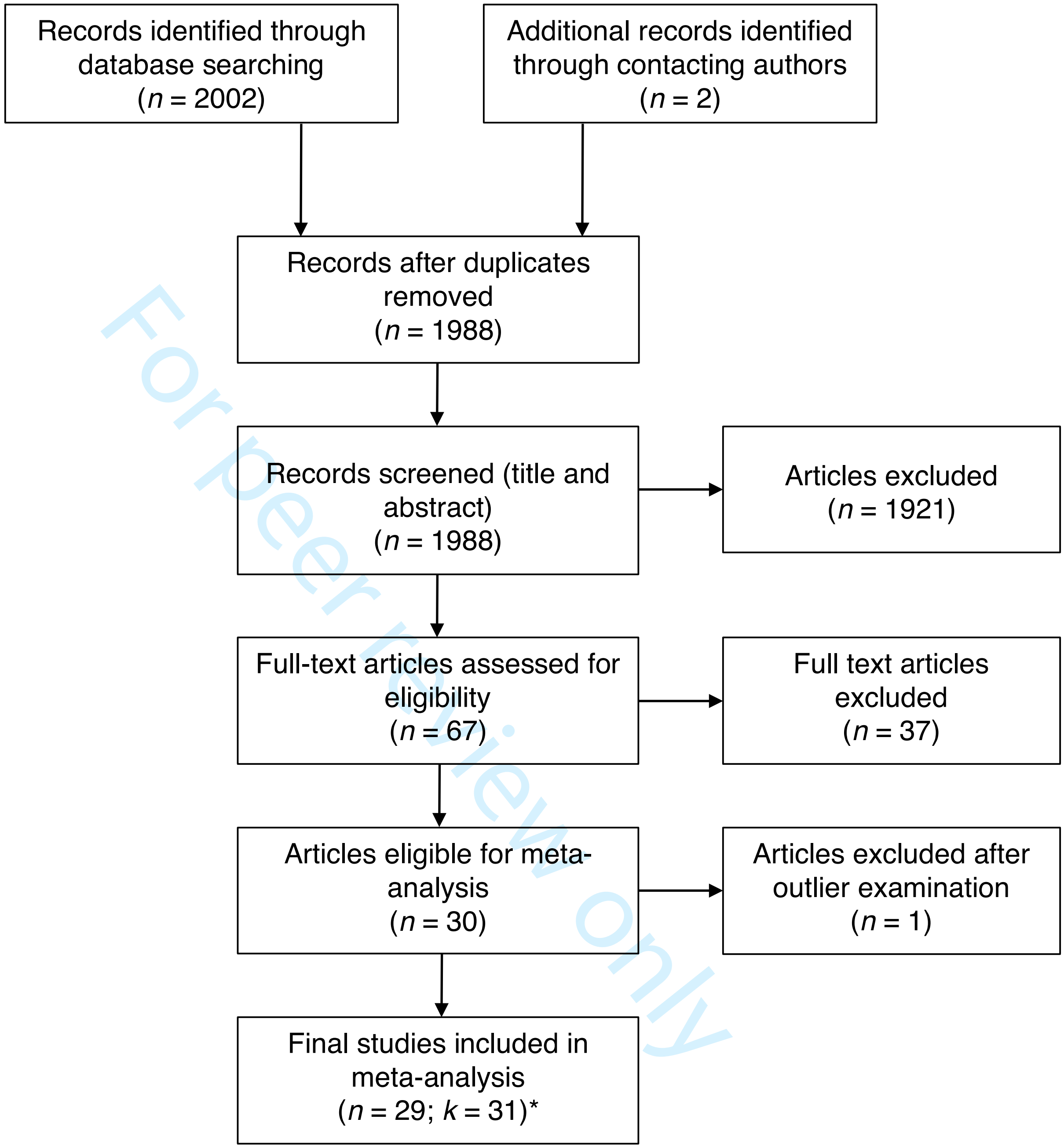
*Note*. $k$ = number of studies; $N$ = cumulative sample size (number of teams); $r$ = sample-size weighted correlation; CI = confidence interval; CR = credibility interval; $Q$ = test statistic for residual heterogeneity of the models; $I^2$ = % of total variability in the effect size estimates due to heterogeneity among true effects (vs. sampling error)

[a] Team size was entered as a continuous variable, therefore, no subgroup analyses exist

## LEGENDS TO FIGURES

**Figure 1** Systematic literature search

**Figure 2** Relationship between teamwork processes and performance


Records identified through database searching (n = 2002)
Additional records identified through contacting authors (n = 2)
Records after duplicates removed (n = 1988)
Records screened (title and abstract) (n = 1988)
Articles excluded (n = 1921)
Full-text articles assessed for eligibility (n = 67)
Full text articles excluded (n = 37)
Articles eligible for meta-analysis (n = 30)
Articles excluded after outlier examination (n = 1)
Final studies included in meta-analysis (n = 29; k = 31)*

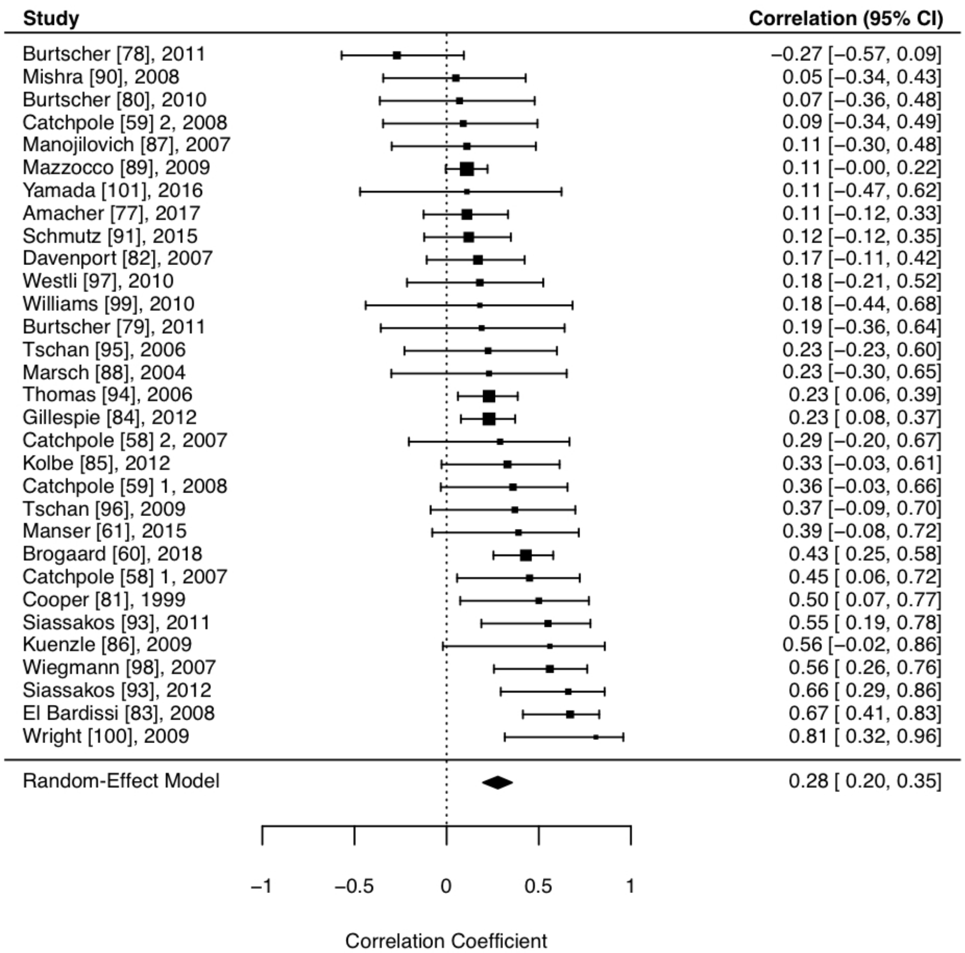



Figure 2. Relationship between teamwork processes and performance.

**SUPPLEMENTARY FILE**

Article: **How effective is teamwork really? The relationship between teamwork and performance in healthcare teams: A systematic review and meta-analysis**

Jan B. Schmutz, PhD, Laurenz L. Meier, PhD, Tanja Manser, PhD

**Search terms used for PubMed search**

(Teamwork[All Fields]) OR (team[All Fields] AND coordination[All Fields]) OR (team[All Fields] AND "decision making"[All Fields]) OR ((team[All Fields]) AND (communication[MeSH Terms]) OR (communication[All Fields])) OR (team[All Fields] AND leadership[All Fields]) AND (patient safety[MeSH Terms])

# PRISMA 2009 Checklist

| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| **TITLE** | | | |
| Title | 1 | Identify the report as a systematic review, meta-analysis, or both. | 1 |
| **ABSTRACT** | | | |
| Structured summary | 2 | Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number. | 2 |
| **INTRODUCTION** | | | |
| Rationale | 3 | Describe the rationale for the review in the context of what is already known. | 4-8 |
| Objectives | 4 | Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS). | 4-5 |
| **METHODS** | | | |
| Protocol and registration | 5 | Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number. | n/a |
| Eligibility criteria | 6 | Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. | 9-10 |
| Information sources | 7 | Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched. | 9 |
| Search | 8 | Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated. | 9 and supplemental material |
| Study selection | 9 | State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). | 9-10 |
| Data collection process | 10 | Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators. | 10 |
| Data items | 11 | List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made. | 9-12 |
| Risk of bias in individual studies | 12 | Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis. | 20 |
| Summary measures | 13 | State the principal summary measures (e.g., risk ratio, difference in means). | 13-19 / 23 |

# PRISMA 2009 Checklist

| | | | |
|---|---|---|---|
| Synthesis of results | 14 | Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $I^2$) for each meta-analysis. | 20-23 |



| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| Risk of bias across studies | 15 | Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies). | 20,26 |
| Additional analyses | 16 | Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified. | 22-23 |
| **RESULTS** | | | |
| Study selection | 17 | Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram. | 21 |
| Study characteristics | 18 | For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations. | 18-19 |
| Risk of bias within studies | 19 | Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12). | 18-19, 20 |
| Results of individual studies | 20 | For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot. | 13-19 |
| Synthesis of results | 21 | Present results of each meta-analysis done, including confidence intervals and measures of consistency. | 23 |
| Risk of bias across studies | 22 | Present results of any assessment of risk of bias across studies (see Item 15). | 23 |
| Additional analysis | 23 | Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]). | N/A |
| **DISCUSSION** | | | |
| Summary of evidence | 24 | Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers). | 24-27 |
| Limitations | 25 | Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias). | 26-27 |
| Conclusions | 26 | Provide a general interpretation of the results in the context of other evidence, and implications for future research. | 27 |
| **FUNDING** | | | |
| Funding | 27 | Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. | 28 |

# PRISMA 2009 Checklist

*From:* Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Page 2 of 2

BMJ Open

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

| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2018-028280.R2 |
| Article Type: | Research |
| Date Submitted by the Author: | 23-Jun-2019 |
| Complete List of Authors: | Schmutz, Jan B.; Northwestern University, Department of Communication Studies<br>Meier, Laurenz ; University of Neuchâtel , Occupational Psychology and Organizations<br>Manser, Tanja; University of Applied Sciences and Arts Northwestern Switzerland |
| <b>Primary Subject Heading</b>: | Communication |
| Secondary Subject Heading: | Communication |
| Keywords: | teamwork, non-technical skills, communication, meta-analysis, teams |
| | |

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

**Jan B. Schmutz, PhD**

Researcher at the Department of Communication Studies, Northwestern University, Evanston, USA

Postdoctoral Researcher and Lecturer at the Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland

jan.schmutz@northwestern.edu

**Laurenz L. Meier, PhD**

Professor at the Institute of Work and Organizational Psychology, University of Neuchâtel, Neuchâtel, Switzerland

laurenz.meier@unine.ch

**Tanja Manser, PhD**

Director of the FHNW School of Applied Psychology, University of Applied Sciences and Arts

Northwestern Switzerland, Olten, Switzerland

tanja.manser@fhnw.ch



Word count: 4770

Keywords: teamwork, non-technical skills, communication, teams, meta-analysis

**Correspondence to**

Jan B. Schmutz
ATLAS Lab
Department of Communication Studies
Northwestern University
Frances Searle Building 1.220
2240 Campus Drive
Evanston, IL 60208
United States
Phone 847-491-7023
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 e-mail.

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

**Data Extraction and Synthesis** Moderator variables (i.e. 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 PRISMA 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 $r$ = .28 (corresponding to an odds ratio 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 organizations should recognize the value of teamwork and emphasize approaches that maintain and improve teamwork for the benefit of their patients.

## ARTICLE SUMMARY

### Strengths and limitations of this study

- This systematic review evaluates available studies investigating the effectiveness of teamwork processes.
- 31 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.

## 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 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 systems[1-4] 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 (e.g. 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 organization 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 (e.g. 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 organizations.

In the following we will first establish an operational definition of teamwork, elaborate on relevant contextual factors, and present our respective meta-analytic 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, *e)* distributed expertise and *f)* clearly assigned roles and responsibilities.[20,21] Based on this definition it becomes clear that teams must dynamically share information and resources amongst 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 (e.g. 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 (i.e. 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* (i.e. 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 emphasizes 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 criticized because of its simplicity.[30] However, it is still the most popular

framework to date and helps to systematize 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 hypothesize 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 (i.e. professional composition, team familiarity, team size); (b) task type (i.e. routine vs. non-routine tasks); (c) two methodological factors related to patient realism (i.e. simulated vs. real) and the type of performance measures used (i.e. 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 to 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 to smaller teams.

*Task type.* Routine situations are characterized by repetitive and unvarying actions (e.g. 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 standardized 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 generalize 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- and outcome-related aspects of performance.[37,38] Process performance measures are action-related aspects and refer to adequate behaviour during procedures (e.g. adhering to guidelines), making them easier to assess. Outcome performance measures (e.g. 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 (e.g. 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 (i.e. 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 (i.e. high process performance) the patient can still die (i.e. 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 PRISMA statement[40] 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 data base PubMed (no limit was placed on the date of publication, last search 19th of June 2018) using the key words *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 8 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 e-mail 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 (e.g. coordination, communication). In addition, studies needed to investigate the relationship between at least one teamwork process and a performance measure (e.g. patient outcome). When studies reported multiple estimates of the same relationship from

the same sample (e.g. 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 (e.g. team cohesion) that influence how a specific team works together.[24] All articles included in this meta-analysis are listed in Table 1 and Table 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-analytic 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. 20% of the studies were rated by both coders. Intercoder 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 (e.g. 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 behavior often under time-pressure (e.g. emergency situations). “Routine tasks” describe previously planned standard procedures (e.g. standard anesthesia 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 (e.g. infection rate, mortality). We focused only on patient-related outcomes and not on team outcomes (e.g. team satisfaction). "Process performance" describes the evaluation of the treatment process and describes how well the process was executed (e.g. 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 (e.g. Odds ratio, $F$ 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. 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 Alpha) or were not available at all for the majority of studies. Therefore, we report uncorrected, sample-size weighted mean correlation, its 95% confidence interval (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.[56] 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"), 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.[56] These models were calculated by the restricted maximum-likelihood estimator, an efficient and unbiased estimator.[57] 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.[58] The results indicate that there is no asymmetry in the funnel plot ($z = 1.79$, $p = .074$), 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.[57]

**Patient and public involvement**

Patients and public were not involved in this study.

## 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.[59,60] After duplicates were removed 1988 articles were screened using title and abstract. 67 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[61,62]). This led to a total of 32 studies coming from 30 articles. Following the recommendation by Viechtbauer and Cheung,

[63] we screened for outliers using studentized deleted residuals. One case (Carlson et al.,[9] $r$ = .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 (e.g. 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 (e.g. time until intervention). Only four studies assessed outcome performance measures. Measures included accuracy of diagnosis, postoperative 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 .28 (95% CI: .20 to .35, $z = 6.55$, $p < .001$), indicating that teamwork is positively related to clinical performance. Results further indicated heterogeneous effect size distributions across the included samples ($Q = 53.73$, $p < .05$, $I^2 = 45.96$), 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 familiarity, team size, task type, patient realism* and *performance measures.*

The omnibus test of moderators was not significant ($F = 0.18$, $df_1 = 6$, $df_2 = 18$, $p > .20$), 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*=.28; [64,65]) 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*=.28 might not seem very high. However, we would like to highlight that *r*=.28 is considered a medium sized effect[64,65] and should not be underestimated. To better illustrate what this effect means we transformed the correlation into an odds ratio (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 (e.g. mortality, morbidity) or are closely related to patient outcomes (e.g. 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.[66] 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 rigor. Guidelines on how to develop performance assessment tools for specific clinical scenarios exist and need to be accounted for.[52,67,68] 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 (e.g. emergencies).[69] Our results suggest that team trainings should not only focus on non-routine situations like emergencies but also on routine situations (e.g. 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 generalize 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 (e.g. 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.[70] A meta-analysis including teams from multiple industries (e.g. 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.[71] Most studies included in our analysis focused on rather short and intense patient care episodes (e.g. 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 data base 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 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[24,26,28,72] 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 to 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.[73,74]

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.[75] 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[76,77] as well as team trainings[78] 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.

**Funding statement:** 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:** All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work; no other relationships or activities that could appear to have influenced the submitted work.

**Author contributions:** 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.

**Ethics approval:** Not required

**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 proof reading the manuscript.

**Data sharing:** Dataset is available from the corresponding author.

## REFERENCE

1 Thomas EJ. Improving teamwork in healthcare: current approaches and the path forward. BMJ Quality & Safety 2011;20:647–50.

2 Salas E, Rosen MA. Building high reliability teams: progress and some reflections on teamwork training. BMJ Quality & Safety 2013;22:369–73.

3 Schraagen JM, Schouten T, Smit M, Haas F, van der Beek D, van de Ven J, et al. A prospective study of paediatric cardiac surgical microsystems: assessing the relationships between non-routine events, teamwork and patient outcomes. BMJ Quality & Safety 2011;20:599–603.

4 Williams AL, Lasky RE, Dannemiller JL, Andrei AM, Thomas EJ. Teamwork behaviours and errors during neonatal resuscitation. BMJ Quality & Safety 2010;19:60–4.

5 Guerlain S, Turrentine FE, Bauer DT, Calland JF, Adams R. Crew resource management training for surgeons: Feasibility and impact. Cognition, Technology & Work 2008;10:255–64.

6 McCulloch P, Mishra A, Handa A, Dale T, Hirst G, Catchpole K. 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, Youngson G, Mitchell A, Rowley D, et al. Debriefing surgeons on non-technical skills (NOTSS). Cogn Tech 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, Manser T. Effective coordination in medical emergency teams: The moderating role of task type. European Journal of Work and Organizational Psychology 2015;24:761–76.

13 Tschan F, Semmer NK, Nägele C, Gurtner A. 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 Castelao E, Russo SG, Riethmüller M, Boos M. 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, Mendez-Tellez PA, Dwyer C, Wyskiel R, et al. A systematic review of behavioural marker systems in healthcare: what do we know about their attributes, validity and application? BMJ Quality & Safety 2014;23:1031–9.

17 Flowerdew L, Brown R, Vincent C, Woloshynowych M. 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, Sonesh SC, Marlow SL, Lacerenza CN, et al. Saving Lives: A Meta-Analysis of Team Training in Healthcare. J Appl Psychol 2016:1266–304.

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. Theoretical Issues in Ergonomics Science 2007;8:381–94.

21 Salas E, Stagl KC, Burke CS, Goodwin GF. Fostering team effectiveness in organizations: Toward an integrative theoretical framework. vol. 52, 2007, p. 185.

22 Fernandez R, Kozlowski SW, Shapiro MJ, Salas E. 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, editors. Team Performance Assessment and Measurement: Theory, methods, and applications, New York: Team performance assessment and measurement: Theory, methods, and applications; 1997, pp. 85–108.

24 Marks MA, Mathieu JE, Zaccaro SJ. A temporally based framework and taxonomy of team processes. The Academy of Management Review 2001;26:356–76.

25 Dietz AS, Pronovost PJ, Mendez-Tellez PA, Wyskiel R, Marsteller JA, Thompson DA, 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, Cuthbertson BH. Developing a team performance framework for the intensive care unit. Crit Care Med 2009;37:1787–93.

27 Ilgen DR, Hollenbeck JR, Johnson M, Jundt D. 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, Pierce L, Kendall D. Understanding team adaptation: a conceptual analysis and model. J Appl Psychol 2006;91:1189–207.

29 Lemieux-Charles L. 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, Beaver GR. Contextualization and Context Theorizing in Teams Research: A Look Back and a Path Forward. The 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 & Social Work 1999;24:210–9.

34 Hall P. Interprofessional teamwork: Professional cultures as barriers. Journal of Interprofessional Care 2005;19:188–96.

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, editors. Essential Simulation in Clinical Education, Oxford, UK: John Wiley & Sons, Ltd; 2013, pp. 213–34.

37 Anderson N, Ones DS, Sinangil HK, Viswesvaran C. Handbook of industrial, work and organizational psychology: Personnel psychology. vol. 1. London: Sage Publications Ltd; 2001.

38 Sonnentag S, Frese M. Performance concepts and performance theory. 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, Mulrow C, Gøtzsche PC, Ioannidis JPA, 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 FL. Methods of meta-analysis: Correcting error and bias in research findings. Thousand Oaks, 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, Glavin RJ, Maran NJ. 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. Importance of teamwork, communication and culture on failure-to-rescue in the elderly. British Journal of Surgery 2016;103.

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, Alves C, Fragata I, Fragata J. 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. Journal of Pediatric Surgery 2016;51:226–30.

48 Kianfar S, Carayon P, Hundt AS, Hoonakker P. Care coordination for chronically ill patients: Identifying coordination activities and interdependencies. Applied Ergonomics. 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. Journal of Applied Psychology 1990;75:334.

51 Campbell JP. Modeling the performance prediction problem in industrial and organizational psychology. In: Dunnette MD, Hough LM, editors. Handbook of industrial and organizational psychology, Palo Alto, CA: Consulting Psychologists Press; 1990, pp. 687–732.

52 Schmutz J, Eppich WJ, Hoffmann F, Heimberg E, Manser T. 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, Rothstein HR. 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 Appl Stat Meth 2003;2:525–30.

55 Lipsey MW, Wilson D. Practical meta-analysis. New York: Sage Publications, Inc;

2001.

56 Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw 2010;36:1–48.

57 Viechtbauer W. Bias and efficiency of meta-analytic variance estimators in the random-effects model. Journal of Educational and Behavioral Statistics 2005;30:261–93.

58 Sterne JA, Egger M. Regression methods to detect publication and other bias in meta-analysis. Publication bias in meta-analysis: Prevention, assessment and adjustments, Chichester: John Wiley & Sons; 2006, pp. 99–110.

59 Brogaard L, Kierkegaard O, Hvidmand L, Jensen KR, Uldbjerg N, Musaeus P, et al. Is non-technical performance the key to high clinical performance in obstetric teams? th Annual conference of the British Maternal Fetal Medicine Society, 2018.

60 Manser T, Bogdanovic J, Clack L., Arora S, Sevdalis N., Schmutz, J. Surgeon's team management skills predict clinical performance. Paper presented at the Annual Meeting of the Society in Europe for Simulation Applied to Medecine (SESAM), Lisbon, Portugal, 2015.

61 Catchpole KR, Giddings AE, Wilkinson M, Hirst G, Dale T, de Leval MR. Improving patient safety by identifying latent failures in successful operations. Surgery 2007;142:102–10.

62 Catchpole K, Mishra A, Handa A, McCulloch P. Teamwork and Error in the Operating Room. Ann Surg 2008;247:699–706.

63 Viechtbauer W, Cheung, MWL. Outlier and influence diagnostics for meta-analysis. Research Synthesis Methods 2010;1:112–125.

64 Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum;

1988.

65 Bosco FA, Aguinis H, Singh K, Field JG, Pierce CA. Correlational effect size benchmarks. Journal of Applied Psychology 2015;100:431.

66 Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat A-HS, Dellinger EP, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. N Engl J Med 2009;360:491–9.

67 Gaba DM, Howard SK, Flanagan B, Smith BE, Fish KJ, Botney R. Assessment of Clinical Performance during Simulated Crises Using Both Technical and Behavioral Ratings. Anesthesiology 1998;89:8–18.

68 Rosen MA, Pronovost PJ. Advancing the use of checklists for evaluating performance in health care. Acad Med 2014;89.

69 Gaba DM, Howard SK, Fish KJ, Smith BE, Sowb YA. Simulation-based training in anesthesia crisis resource management (ACRM): a decade of experience. Simulation & Gaming 2001;32:175–93.

70 Wageman R, Baker G. Incentives and cooperation: The joint effects of task and reward interdependence on group performance. J Organiz Behav 1997:139–58.

71 LePine JA, Piccolo RF, Jackson CL, Mathieu JE, Saul JR. A meta-analysis of teamwork processes: tests of a multidimensional model and relationships with team effectiveness criteria. Pers Psychol 2008;61:273–307.

72 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. European Journal of Work and Organizational Psychology 2015;24:1–26.

73 Edmondson AC. Teaming: How organizations learn, innovate, and compete in the

knowledge economy. San Francisco, CA: Jossey-Bass; 2012.

74 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.

75 Henrickson Parker, S., Schmutz, J. B., & Manser, T. Training needs for adaptive coordination: Utilizing task analysis to identify coordination requirements in three different clinical settings. Gr & Org Manag 2018;43:504-527.

76 Eppich WJ, Mullan PC, Brett-Fleegler M, Cheng A. “Let's Talk About It”: Translating Lessons From Health Care Simulation to Clinical Event Debriefings and Coaching Conversations. Clinical Pediatric Emergency Medicine 2016;17:200–11.

77 Schmutz JB, Eppich WJ. Promoting learning and patient care through shared reflection: A conceptual framework for team reflexivity in healthcare. Acad Med 2017;92:1555–63.

78 Salas E, DiazGranados D, Weaver SJ, King H. Does Team Training Work? Principles for Health Care. Acad Emerg Med 2008;15:1002–9.

79 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.

80 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.

81 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.

82 Burtscher MJ, Wacker J, Grote G, et al. Managing nonroutine events in anesthesia: the role of adaptive coordination. *Hum Factors* 2010;52:282–94.

83 Cooper S, Wakelam A. Leadership of resuscitation teams: "Lighthouse Leadership." *Resuscitation* 1999;42:27–45.

84 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.

85 El Bardissi 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.

86 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.

87 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.

88 Künzle B, Zala-Mezo E, Wacker J. Leadership in anaesthesia teams: the most effective leadership is shared. *Qual Saf Health Care* 2010;19:1–6.

89 Manojlovich M, DeCicco B. Healthy Work Environments, Nurse-Physician Communication, and Patients Outcomes. *Am J Crit Care* 2007;16:536–43.

90 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.

91 Mazzocco K, Petitti DB, Fong KT, et al. Surgical team behaviors and patient outcomes.

*Am J Surg* 2009;197:678–85.

92 Mishra A, Catchpole K, Dale T, et al. The influence of non-technical performance on technical outcome in laparoscopic cholecystectomy. *Surg Endosc* 2007;22:68–73.

93 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.

94 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;32:41.

95 Siassakos D, Fox R, Crofts JF, et al. The management of a simulated emergency: Better teamwork, better performance. *Resuscitation*

96 Thomas EJ, Sexton JB, Lasky RE, et al. Teamwork and quality during neonatal care in the delivery room. *J Perinatol* 2006;26:163–9.

97 Tschan F, Semmer NK, Gautschi, et al. Leading to Recovery: Group Performance and Coordinative Activities in Medical Emergency Driven Groups. *Hum Perf* 2006;19:277–304.

98 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.

99 Westli H, Johnsen B, 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.

100 Wiegmann DA, ElBardissi AW, Dearani JA. Disruptions in surgical flow and their

relationship to surgical errors: an exploratory investigation. *Surgery* 2007;142:658–65.

101 Williams AL, Lasky RE, Dannemiller JL, et al. Teamwork behaviours and errors during neonatal resuscitation. *Qual Saf Health Care* 2010;19:60–4.

102 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.

103 Yamada NK, Fuerch JH, Halamek LP. Impact of standardized communication techniques on errors during simulated neonatal resuscitation. *AM J Perinatol* 2016;33:385–92.

*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, Schumacher, Legeret, et al.[79] | 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<br>Hands-on time within first 180 seconds |
| Brogaard, Kierkegaard, Hvidmand, et al.[59] | 2018 | To investigate the relationship between non-technical skills and clinical performance in obstetric teams | Video observation of obstetric teams (Obstetricians, obstetric 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, Kolbe, Wacker, et al.[80] | 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, Manser, Kolbe, et al.[81] | 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, Wacker, Grote, et al.[82] | 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, Min, Bridges[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, Giddings, Wilkinson, et al.[61] | 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, Mishra, Handa, et al.[62] | 2008 | To analyse the effects of surgical, aesthetic, 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 |
| Cooper, Wakelam[83] | 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 minutes | Survey about leadership behaviour using the Leadership Behaviour Description Questionnaire | Checklist based expert rating |
| Davenport, Henderson, Mosca, et al.[84] | 2007 | To measure the impact of organizational 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<br>Surgical mortality |
| El Bardissi, Wiegmann, Henrickson, et al.[85] | 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, Chaboyer, Fairweather[86] | 2012 | To investigate how various human factors variables, extend the expected length of an operation | Live observation of surgical teams across 10 specialties | Structured observation of numbers of communication failures | Deviation from expected length of operation |

| | | | | | |
|---|---|---|---|---|---|
| Kolbe, Burtscher, Wacker, et al.[87] | 2012 | To test the relationship between speaking up and technical team performance in anaesthesia. | Observation of 2-person (nurse, resident) ad hoc anaesthesia teams performing simulated inductions of general anaesthesia with minor nonroutine events | Structured observation of speaking up behaviour | Checklist based expert rating |
| Kuenzle, Zala-Mezo, Wacker, et al.[88] | 2009 | To investigate shared leadership patterns during anaesthesia induction and to show how they are linked to team performance | Observation of 2-person (nurse, resident) ad hoc anaesthesia teams performing simulated inductions of general anaesthesia with a nonroutine event (asystole) | Structured observation of leadership behaviour | Reaction time to nonroutine event |
| Manojilovich, Antonakos, David, et al.[89] | 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<br>Bloodstream infections<br>Pressure ulcers<br>Acute physiology and chronic health evaluation score |
| Manser, Bogdanovic, Clack, et al. [60] | 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, Müller, Marquardt, et al.[90] | 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 |
| Mazzocco, Petitti, Fong, et al.[91] | 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, Catchpole, Dale, et al.[92] | 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, Hoffmann, Heimberg, et al.[93] | 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, Bristowe, Draycott, et al.[94] | 2012 | To investigate the relationship between patient satisfaction and communication | Video observation of teams (physicians, midwives) managing obstetric emergencies in secondary and tertiary maternity units | Structured observation of closed loop communication | Timely administration of magnesium sulphate |
| Siassakos, Fox, Crofts, et al.[95] | 2011 | To determine whether team performance in a simulated emergency is related to generic teamwork skills and behaviors | 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, Sexton, Lasky, et al.[96] | 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 | Compliance with Neonatal Resuscitation Program guidelines |
| Tschan, Semmer, Gautschi, et al.[97] | 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, Semmer, Gurtner, et al.[98] | 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 |
| Westli, Johnsen, Eid, et al.[99] | 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, El Bardissi, Dearani, et al.[100] | 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, Lasky, Dannemiller, et al.[101] | 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, Phillips-Bute, Petrusa, et al.[102] | 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, Fuerch, Halamek[103] | 2016 | To investigate the effect of standardized 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<br>Time to initiate positive pressure ventilation<br>Time to chest compression |

*Table 2.* Studies, effect sizes and moderator variables included in the meta-analytic database

| Authors | Year | Study goal | Setting | No. of teams | Professional composition | Team famil-iarity | Average team size | Task type | Patient realism | Perfor-mance measure |
|---|---|---|---|---|---|---|---|---|---|---|
| Amacher, Schumacher, Legeret, et al.[79] | 2017 | .11 | Emergency medicine | 72 | Uniprofessional | Experi-mental | 3 | Non-routine | Simulated | Process |
| Brogaard, Kierkegaard, Hvidmand, et al.[59] | 2018 | .43 | Obstetrics | 99 | Interprofessional | Real | 5 | Non-routine | Real | Process |
| Burtscher, Kolbe, Wacker, et al.[80] | 2011 | -.27 | Anaesthesia | 31 | Interprofessional | Experi-mental | 2 | Routine | Simulated | Process |
| Burtscher, Manser, Kolbe, et al.[81] | 2011 | .19 | Anaesthesia | 15 | Interprofessional | Experi-mental | 2 | Routine & non-routine | Simulated | Process |
| Burtscher, Wacker, Grote, et al.[82] | 2010 | .07 | Anaesthesia | 22 | Interprofessional | Real | 3 | Non-routine | Real | Process |
| Carlson, Min, Bridges[9][b] | 2009 | .83 | Emergency medicine | 44 | Uniprofessional | Experi-mental | 2.6 | Non-routine | Simulated | Process |
| Catchpole, Giddings, Wilkinson, et al.[61] | 2007 | .45† | Surgery | 24 | Interprofessional | Real | 9 | Non-routine | Real | Process |
| | 2007 | .29† | Surgery | 18 | Interprofessional | Real | 5 | Routine | Real | Process |
| Catchpole, Mishra, Handa, et al.[62] | 2008 | .36† | Surgery | 26 | Interprofessional | Real | | Routine | Real | Process |
| | 2008 | .09† | Surgery | 22 | Interprofessional | Real | | Routine | Real | Process |
| Cooper, Wakelam[83] | 1999 | .50 | General care | 20 | Interprofessional | Real | 4 | Routine | Real | Process |
| Davenport, Henderson, Mosca, et al.[84] | 2007 | .17 | Surgery | 52 | Interprofessional | Real | | Routine | Real | Outcome |
| El Bardissi, Wiegmann, Henrickson, et al.[85] | 2008 | .67 | Surgery | 31 | Interprofessional | Real | 7 | Routine | Real | Process |
| Gillespie, Chaboyer, Fairweather[86] | 2012 | .23 | Surgery | 160 | Interprofessional | Real | 6 | Routine | Real | Process |
| Kolbe, Burtscher, Wacker, et al.[87] | 2012 | .33 | Anaesthesia | 31 | Interprofessional | Real | 2 | Non-routine | Simulated | Process |
| Kuenzle, Zala-Mezo, Wacker, et al.[88] | 2009 | .56 | Anaesthesia | 12 | Interprofessional | Real | 2 | Routine | Simulated | Process |
| Manojilovich, Antonakos, David, et al.[89] | 2009 | .11 | Intensive care | 25 | Uniprofessional | Real | 36 | Routine | Real | Outcome |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Manser, Bogdanovic, Clack, et al.[60] | 2015 | .39 | Surgery | 19 | Interprofessional | Experi-mental | 5 | Routine | Simulated | Process |
| Marsch, Müller, Marquardt, et al.[90] | 2004 | .23 | Intensive care | 16 | Interprofessional | Experi-mental | 3 | Non-routine | Simulated | Process |
| Mazzocco, Petitti, Fong, et al.[91] | 2009 | .11 | Surgery | 293 | Interprofessional | Real | 6 | Routine | Real | Outcome |
| Mishra, Catchpole, Dale, et al.[92] | 2008 | .05 | Surgery | 26 | Interprofessional | Real | 6 | Routine | Real | Process |
| Schmutz, Hoffmann, Heimberg, et al.[93] | 2015 | .12 | Emergency medicine | 68 | Interprofessional | Real | 6 | Non-routine | Simulated | Process |
| Siassakos, Bristowe, Draycott, et al.[94] | 2012 | .66 | Obstetrics | 19 | Interprofessional | Real | 6 | Non-routine | Simulated | Process |
| Siassakos, Fox, Crofts, et al.[95] | 2011 | .55 | Emergency medicine/ obstetrics | 24 | Interprofessional | Experi-mental | 6 | Non-routine | Simulated | Process |
| Thomas, Sexton, Lasky, et al.[96] | 2006 | .23 | Neonatal care | 132 | Interprofessional | Real | 5 | Non-routine | Real | Process |
| Tschan, Semmer, Gautschi, et al.[97] | 2006 | .23 | Emergency medicine | 21 | Interprofessional | Experi-mental | 5 | Non-routine | Simulated | Process |
| Tschan, Semmer, Gurtner, et al.[98] | 2009 | .37 | Emergency medicine | 20 | Uniprofessional | Experi-mental | 2.65 | Non-routine | Simulated | Outcome |
| Westli, Johnsen, Eid, et al.[99] | 2010 | .18 | Emergency medicine | 27 | Interprofessional | Real | 5.1 | Non-routine | Simulated | Process |
| Wiegmann, El Bardissi, Dearani, et al.[100] | 2007 | .56 | Surgery | 31 | Interprofessional | Real | | Routine | Real | Process |
| Williams, Lasky, Dannemiller, et al.[101] | 2010 | .18 | Neonatal care | 12 | Interprofessional | Real | 5 | Non-routine | Real | Process |
| Wright, Phillips-Bute, Petrusa, et al.[102] | 2009 | .81 | General care | 9 | Uniprofessional | Experi-mental | 4 | Non-routine | Simulated | Process |
| Yamada, Fuerch, Halamek[103] | 2016 | .11 | Emergency medicine | 13 | Interprofessional | Experi-mental | 3 | Non-routine | Simulated | Process |

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

[b] Carlson, Min & Bridges has been identified as an outlier and therefore excluded from the analysis

*Table 3.* Meta-Analytic relationships between teamwork and clinical performance

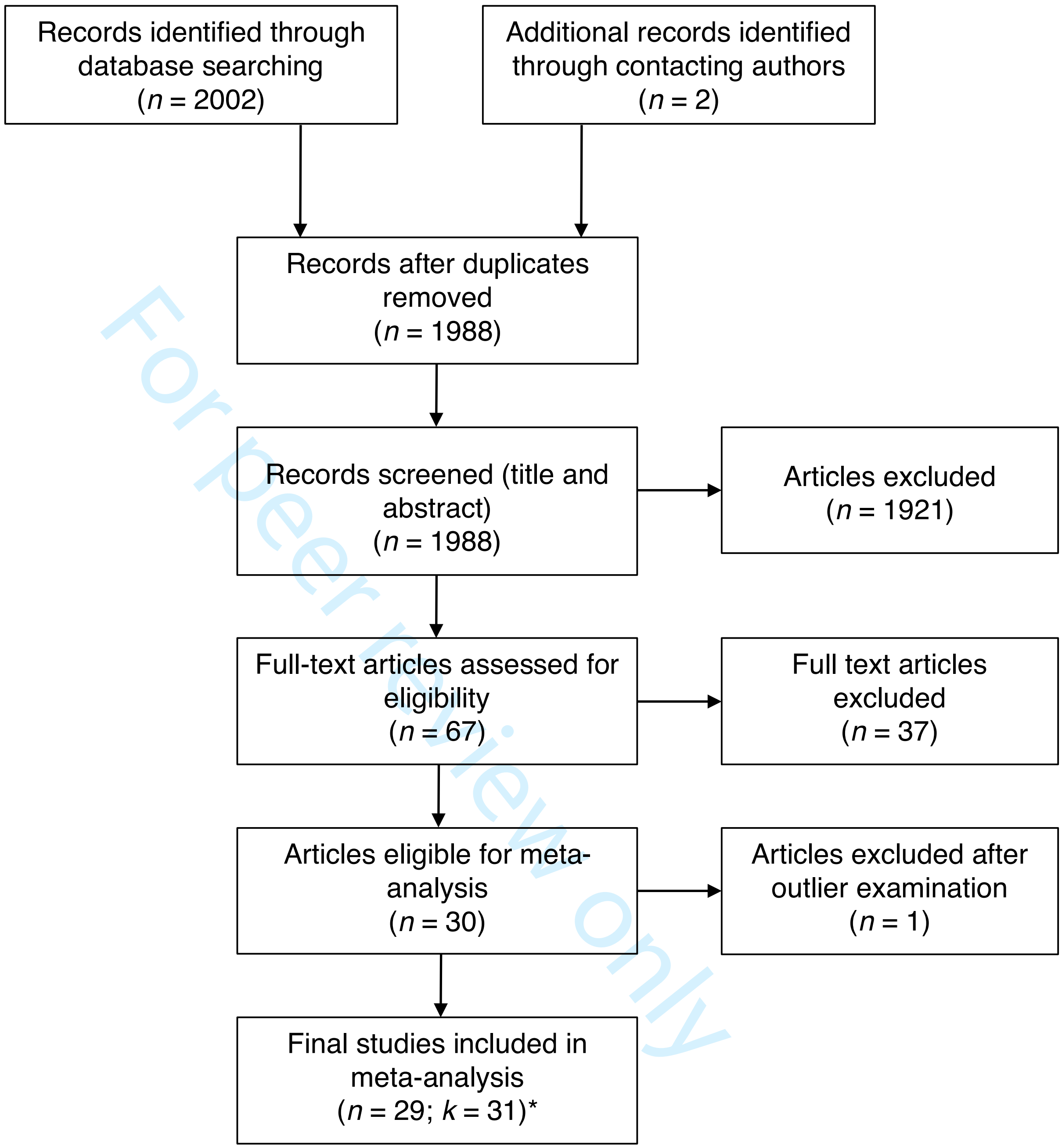
| | N | k | r | 95% CI | 80% CR | Q | $I^2$ |
|---|---|---|---|---|---|---|---|
| Overall relationship | 1,390 | 31 | .28* | [.20 ; .35] | [.09 ; .45] | 53.7* | 46.0 |
| *Team characteristics* | | | | | | | |
| Professional composition | | | | | | | |
| Interprofessional | 1,264 | 27 | .28* | [.20 ; .36] | [.09 ; .46] | 47.1* | 48.2 |
| Uniprofessional | 126 | 4 | .28 | [-.01 ; .52] | [-.04 ; .54] | 6.5 | 47.1 |
| Team familiarity | | | | | | | |
| Experimental team | 240 | 10 | .25* | [.05 ; .43] | [-.05 ; .51] | 17.2* | 47.2 |
| Real team | 1,150 | 21 | .29* | [.20 ; .37] | [.12 ; .45] | 36.2* | 45.7 |
| Team size[a] | | | | | | | |
| *Task characteristics* | | | | | | | |
| Task type | | | | | | | |
| Routine task | 766 | 14 | .27* | [.12 ; .40] | [-.01 ; .50] | 30.9* | 65.0 |
| Non-routine task | 609 | 16 | .29* | [.20 ; .39] | [.16 ; .42] | 20.5 | 24.6 |
| *Methodological factors* | | | | | | | |
| Patient realism | | | | | | | |
| Real patient | 993 | 16 | .28* | [.18 ; .38] | [.10 ; .45] | 28.7* | 49.3 |
| Simulated patient | 397 | 15 | .28* | [.13 ; .41] | [.02 ; .50] | 25.0* | 44.6 |
| Performance measures | | | | | | | |
| Outcome performance | 390 | 4 | .13* | [.03 ; .23] | [.06 ; .19] | 1.3 | 0.0 |
| Process performance | 1,000 | 27 | .30* | [.21 ; .39] | [.10 ; .49] | 45.6* | 45.6 |

*Note*. $k$ = number of studies; $N$ = cumulative sample size (number of teams); $r$ = sample-size weighted correlation; CI = confidence interval; CR = credibility interval; $Q$ = test statistic for residual heterogeneity of the models; $I^2$ = % of total variability in the effect size estimates due to heterogeneity among true effects (vs. sampling error)
[a] Team size was entered as a continuous variable, therefore, no subgroup analyses exist
* $p < .05$.

## LEGENDS TO FIGURES

**Figure 1** Systematic literature search

**Figure 2** Relationship between teamwork processes and performance


Records identified through database searching (*n* = 2002)
Additional records identified through contacting authors (*n* = 2)
Records after duplicates removed (*n* = 1988)
Records screened (title and abstract) (*n* = 1988)
Articles excluded (*n* = 1921)
Full-text articles assessed for eligibility (*n* = 67)
Full text articles excluded (*n* = 37)
Articles eligible for meta-analysis (*n* = 30)
Articles excluded after outlier examination (*n* = 1)
Final studies included in meta-analysis (*n* = 29; *k* = 31)*

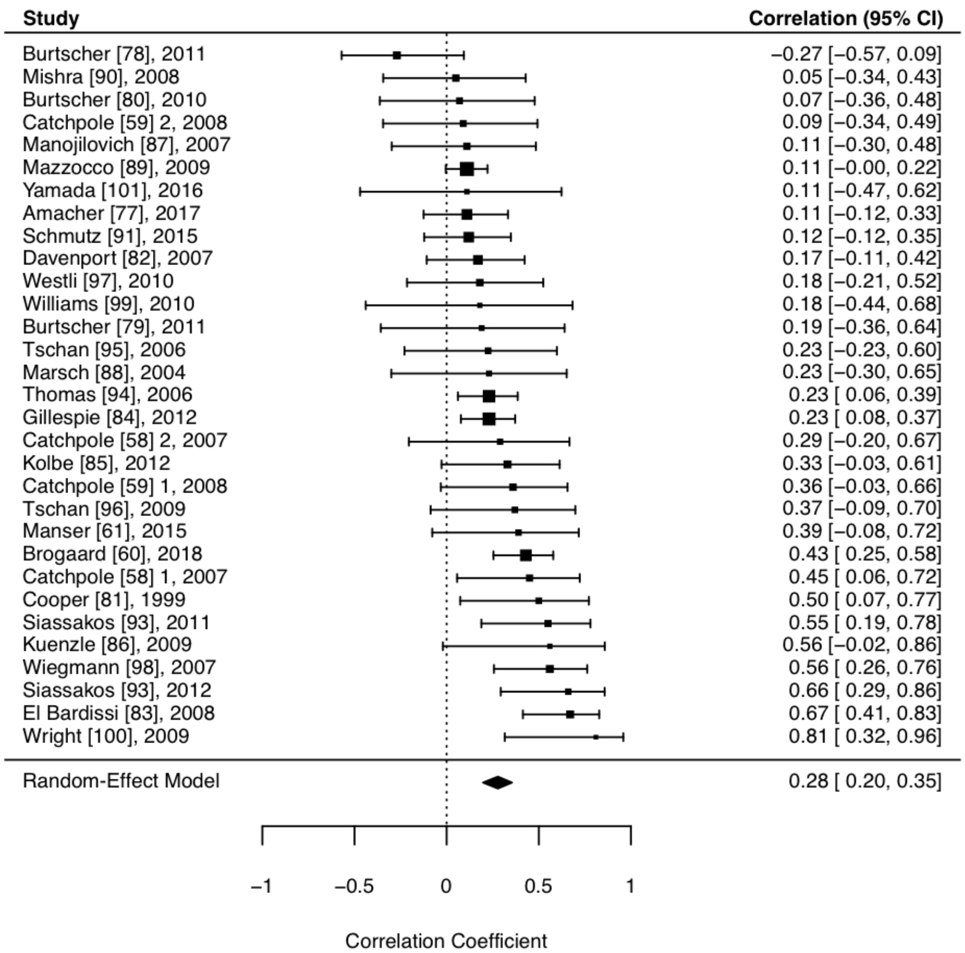



Figure 2. Relationship between teamwork processes and performance.

**SUPPLEMENTARY FILE**

Article: **How effective is teamwork really? The relationship between teamwork and performance in healthcare teams: A systematic review and meta-analysis**

Jan B. Schmutz, PhD, Laurenz L. Meier, PhD, Tanja Manser, PhD

**Search terms used for PubMed search**

(Teamwork[All Fields]) OR (team[All Fields] AND coordination[All Fields]) OR (team[All Fields] AND "decision making"[All Fields]) OR ((team[All Fields]) AND (communication[MeSH Terms]) OR (communication[All Fields])) OR (team[All Fields] AND leadership[All Fields]) AND (patient safety[MeSH Terms])

# PRISMA 2009 Checklist

| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| **TITLE** | | | |
| Title | 1 | Identify the report as a systematic review, meta-analysis, or both. | 1 |
| **ABSTRACT** | | | |
| Structured summary | 2 | Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number. | 2 |
| **INTRODUCTION** | | | |
| Rationale | 3 | Describe the rationale for the review in the context of what is already known. | 4-8 |
| Objectives | 4 | Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS). | 4-5 |
| **METHODS** | | | |
| Protocol and registration | 5 | Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number. | n/a |
| Eligibility criteria | 6 | Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. | 9-10 |
| Information sources | 7 | Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched. | 9 |
| Search | 8 | Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated. | 9 and supplemental material |
| Study selection | 9 | State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). | 9-10 |
| Data collection process | 10 | Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators. | 10 |
| Data items | 11 | List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made. | 9-12 |
| Risk of bias in individual studies | 12 | Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis. | 20 |
| Summary measures | 13 | State the principal summary measures (e.g., risk ratio, difference in means). | 13-19 / 23 |

# PRISMA 2009 Checklist

| | | | |
|---|---|---|---|
| Synthesis of results | 14 | Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $I^2$) for each meta-analysis. | 20-23 |

Page 1 of 2

| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| Risk of bias across studies | 15 | Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies). | 20,26 |
| Additional analyses | 16 | Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified. | 22-23 |
| **RESULTS** | | | |
| Study selection | 17 | Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram. | 21 |
| Study characteristics | 18 | For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations. | 18-19 |
| Risk of bias within studies | 19 | Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12). | 18-19, 20 |
| Results of individual studies | 20 | For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot. | 13-19 |
| Synthesis of results | 21 | Present results of each meta-analysis done, including confidence intervals and measures of consistency. | 23 |
| Risk of bias across studies | 22 | Present results of any assessment of risk of bias across studies (see Item 15). | 23 |
| Additional analysis | 23 | Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]). | N/A |
| **DISCUSSION** | | | |
| Summary of evidence | 24 | Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers). | 24-27 |
| Limitations | 25 | Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias). | 26-27 |
| Conclusions | 26 | Provide a general interpretation of the results in the context of other evidence, and implications for future research. | 27 |
| **FUNDING** | | | |
| Funding | 27 | Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. | 28 |

# PRISMA 2009 Checklist

*From:* Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

BMJ Open

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

| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2018-028280.R3 |
| Article Type: | Research |
| Date Submitted by the Author: | 14-Aug-2019 |
| Complete List of Authors: | Schmutz, Jan B.; Northwestern University, Department of Communication Studies<br>Meier, Laurenz ; University of Neuchâtel , Occupational Psychology and Organizations<br>Manser, Tanja; University of Applied Sciences and Arts Northwestern Switzerland |
| <b>Primary Subject Heading</b>: | Communication |
| Secondary Subject Heading: | Communication |
| Keywords: | teamwork, non-technical skills, communication, meta-analysis, teams |
| | |

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

**Jan B. Schmutz, PhD**

Researcher at the Department of Communication Studies, Northwestern University, Evanston, USA

Postdoctoral Researcher and Lecturer at the Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland

jan.schmutz@northwestern.edu

**Laurenz L. Meier, PhD**

Professor at the Institute of Work and Organizational Psychology, University of Neuchâtel, Neuchâtel, Switzerland

laurenz.meier@unine.ch

**Tanja Manser, PhD**

Director of the FHNW School of Applied Psychology, University of Applied Sciences and Arts

Northwestern Switzerland, Olten, Switzerland

tanja.manser@fhnw.ch



Word count: 4852

Keywords: teamwork, non-technical skills, communication, teams, meta-analysis

**Correspondence to**

Jan B. Schmutz
ATLAS Lab
Department of Communication Studies
Northwestern University
Frances Searle Building 1.220
2240 Campus Drive
Evanston, IL 60208
United States
Phone 847-491-7023
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 e-mail.

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

**Data Extraction and Synthesis** Moderator variables (i.e. 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 PRISMA 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 $r$ = .28 (corresponding to an odds ratio 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 organizations should recognize the value of teamwork and emphasize approaches that maintain and improve teamwork for the benefit of their patients.

## ARTICLE SUMMARY

### Strengths and limitations of this study

- This systematic review evaluates available studies investigating the effectiveness of teamwork processes.
- 31 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.

# 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 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 systems[1-4] 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 (e.g. 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 organization 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 (e.g. 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 organizations.

In the following we will first establish an operational definition of teamwork, elaborate on relevant contextual factors, and present our respective meta-analytic 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, *e)* distributed expertise and *f)* clearly assigned roles and responsibilities.[20,21] Based on this definition it becomes clear that teams must dynamically share information and resources amongst 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 (e.g. 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 (i.e. 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* (i.e. 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 emphasizes 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 criticized because of its simplicity.[30] However, it is still the most popular

framework to date and helps to systematize 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 hypothesize 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 (i.e. professional composition, team familiarity, team size); (b) task type (i.e. routine vs. non-routine tasks); (c) two methodological factors related to patient realism (i.e. simulated vs. real) and the type of performance measures used (i.e. 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 to 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 to smaller teams.

*Task type.* Routine situations are characterized by repetitive and unvarying actions (e.g. 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 standardized 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 generalize 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- and outcome-related aspects of performance.[37,38] Process performance measures are action-related aspects and refer to adequate behaviour during procedures (e.g. adhering to guidelines), making them easier to assess. Outcome performance measures (e.g. 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 (e.g. 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 (i.e. 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 (i.e. high process performance) the patient can still die (i.e. 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 PRISMA statement[40] 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 data base PubMed (no limit was placed on the date of publication, last search 19th of June 2018) using the key words *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 8 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 e-mail 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 (e.g. coordination, communication). In addition, studies needed to investigate the relationship between at least one teamwork process and a performance measure (e.g. patient outcome). When studies reported multiple estimates of the same relationship from

the same sample (e.g. 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 (e.g. team cohesion) that influence how a specific team works together.[24] All articles included in this meta-analysis are listed in Table 1 and Table 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-analytic 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. 20% of the studies were rated by both coders. Intercoder 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 (e.g. 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 behavior often under time-pressure (e.g. emergency situations). “Routine tasks” describe previously planned standard procedures (e.g. standard anesthesia 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 (e.g. infection rate, mortality). We focused only on patient-related outcomes and not on team outcomes (e.g. team satisfaction). "Process performance" describes the evaluation of the treatment process and describes how well the process was executed (e.g. 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 (e.g. Odds ratio, $F$ 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[1]. 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 Alpha) or were not available at all for the majority of studies. Therefore, we report uncorrected, sample-size weighted mean correlation, its 95% confidence interval (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.

[1] Scholars have suggested to convert $r$ to Fisher's z scores, to average the z's, and then to backtransform it to $r$.[56] Using simple arithmetic average (i.e., 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.

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"), 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 ($z = 1.79$, $p = .074$), 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]

**Patient and public involvement**

Patients and public were not involved in this study.

## 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. 67 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] $r$ = .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 (e.g. 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 (e.g. time until intervention). Only four studies assessed outcome performance measures. Measures included accuracy of diagnosis, postoperative 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 .28 (95% CI: .20 to .35, $z = 6.55$, $p < .001$), indicating that teamwork is positively related to clinical performance. Results further indicated heterogeneous effect size distributions across the included samples ($Q = 53.73$, $p < .05$, $I^2 = 45.96$), 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 familiarity, team size, task type, patient realism* and *performance measures.*

The omnibus test of moderators was not significant ($F = 0.18$, $df_1 = 6$, $df_2 = 18$, $p > .20$), 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*=.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*=.28 might not seem very high. However, we would like to highlight that *r*=.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 odds ratio (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 (e.g. mortality, morbidity) or are closely related to patient outcomes (e.g. 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 rigor. 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 (e.g. emergencies).[70] Our results suggest that team trainings should not only focus on non-routine situations like emergencies but also on routine situations (e.g. 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 generalize 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 (e.g. 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 (e.g. 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 (e.g. 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 data base 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 to 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 trainings[79] 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.

**Funding statement:** 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:** All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work; no other relationships or activities that could appear to have influenced the submitted work.

**Author contributions:** 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.

**Ethics approval:** Not required

**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 proof reading the manuscript.

**Data sharing:** Dataset is available from the corresponding author.

**REFERENCE**

1 Thomas EJ. Improving teamwork in healthcare: current approaches and the path forward. BMJ Quality & Safety 2011;20:647–50.

2 Salas E, Rosen MA. Building high reliability teams: progress and some reflections on teamwork training. BMJ Quality & Safety 2013;22:369–73.

3 Schraagen JM, Schouten T, Smit M, Haas F, van der Beek D, van de Ven J, et al. A prospective study of paediatric cardiac surgical microsystems: assessing the relationships between non-routine events, teamwork and patient outcomes. BMJ Quality & Safety 2011;20:599–603.

4 Williams AL, Lasky RE, Dannemiller JL, Andrei AM, Thomas EJ. Teamwork behaviours and errors during neonatal resuscitation. BMJ Quality & Safety 2010;19:60–4.

5 Guerlain S, Turrentine FE, Bauer DT, Calland JF, Adams R. Crew resource management training for surgeons: Feasibility and impact. Cognition, Technology & Work 2008;10:255–64.

6 McCulloch P, Mishra A, Handa A, Dale T, Hirst G, Catchpole K. 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, Youngson G, Mitchell A, Rowley D, et al. Debriefing surgeons on non-technical skills (NOTSS). Cogn Tech 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, Manser T. Effective coordination in medical emergency teams: The moderating role of task type. European Journal of Work and Organizational Psychology 2015;24:761–76.

13 Tschan F, Semmer NK, Nägele C, Gurtner A. 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 Castelao E, Russo SG, Riethmüller M, Boos M. 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, Mendez-Tellez PA, Dwyer C, Wyskiel R, et al. A systematic review of behavioural marker systems in healthcare: what do we know about their attributes, validity and application? BMJ Quality & Safety 2014;23:1031–9.

17 Flowerdew L, Brown R, Vincent C, Woloshynowych M. 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, Sonesh SC, Marlow SL, Lacerenza CN, et al. Saving Lives: A Meta-Analysis of Team Training in Healthcare. J Appl Psychol 2016:1266–304.

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. Theoretical Issues in Ergonomics Science 2007;8:381–94.

21 Salas E, Stagl KC, Burke CS, Goodwin GF. Fostering team effectiveness in organizations: Toward an integrative theoretical framework. vol. 52, 2007, p. 185.

22 Fernandez R, Kozlowski SW, Shapiro MJ, Salas E. 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, editors. Team Performance Assessment and Measurement: Theory, methods, and applications, New York: Team performance assessment and measurement: Theory, methods, and applications; 1997, pp. 85–108.

24 Marks MA, Mathieu JE, Zaccaro SJ. A temporally based framework and taxonomy of team processes. The Academy of Management Review 2001;26:356–76.

25 Dietz AS, Pronovost PJ, Mendez-Tellez PA, Wyskiel R, Marsteller JA, Thompson DA, 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, Cuthbertson BH. Developing a team performance framework for the intensive care unit. Crit Care Med 2009;37:1787–93.

27 Ilgen DR, Hollenbeck JR, Johnson M, Jundt D. 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, Pierce L, Kendall D. Understanding team adaptation: a conceptual analysis and model. J Appl Psychol 2006;91:1189–207.

29 Lemieux-Charles L. 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, Beaver GR. Contextualization and Context Theorizing in Teams Research: A Look Back and a Path Forward. The 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 & Social Work 1999;24:210–9.

34 Hall P. Interprofessional teamwork: Professional cultures as barriers. Journal of Interprofessional Care 2005;19:188–96.

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, editors. Essential Simulation in Clinical Education, Oxford, UK: John Wiley & Sons, Ltd; 2013, pp. 213–34.

37 Anderson N, Ones DS, Sinangil HK, Viswesvaran C. Handbook of industrial, work and organizational psychology: Personnel psychology. vol. 1. London: Sage Publications Ltd; 2001.

38 Sonnentag S, Frese M. Performance concepts and performance theory. 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, Mulrow C, Gøtzsche PC, Ioannidis JPA, 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 FL. Methods of meta-analysis: Correcting error and bias in research findings. Thousand Oaks, 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, Glavin RJ, Maran NJ. 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. Importance of teamwork, communication and culture on failure-to-rescue in the elderly. British Journal of Surgery 2016;103.

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, Alves C, Fragata I, Fragata J. 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. Journal of Pediatric Surgery 2016;51:226–30.

48 Kianfar S, Carayon P, Hundt AS, Hoonakker P. Care coordination for chronically ill patients: Identifying coordination activities and interdependencies. Applied Ergonomics. 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. Journal of Applied Psychology 1990;75:334.

51 Campbell JP. Modeling the performance prediction problem in industrial and organizational psychology. In: Dunnette MD, Hough LM, editors. Handbook of industrial and organizational psychology, Palo Alto, CA: Consulting Psychologists Press; 1990, pp. 687–732.

52 Schmutz J, Eppich WJ, Hoffmann F, Heimberg E, Manser T. 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, Rothstein HR. 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 Appl 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 App Psy*. 1983;68:78.

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. Journal of Educational and Behavioral Statistics 2005;30:261–93.

59 Sterne JA, Egger M. Regression methods to detect publication and other bias in meta-analysis. Publication bias in meta-analysis: Prevention, assessment and adjustments, Chichester: John Wiley & Sons; 2006, pp. 99–110.

60 Brogaard L, Kierkegaard O, Hvidmand L, Jensen KR, Uldbjerg N, Musaeus P, et al. 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., Arora S, Sevdalis N., Schmutz, J. Surgeon's team management skills predict clinical performance. Paper presented at the Annual Meeting of the Society in Europe for Simulation Applied to Medecine (SESAM), Lisbon, Portugal, 2015.

62 Catchpole KR, Giddings AE, Wilkinson M, Hirst G, Dale T, de Leval MR. Improving patient safety by identifying latent failures in successful operations. Surgery 2007;142:102–10.

63 Catchpole K, Mishra A, Handa A, McCulloch P. Teamwork and Error in the Operating Room. Ann Surg 2008;247:699–706.

64 Viechtbauer W, Cheung, MWL. Outlier and influence diagnostics for meta-analysis.

Research Synthesis Methods 2010;1:112–125.

65 Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum; 1988.

66 Bosco FA, Aguinis H, Singh K, Field JG, Pierce CA. Correlational effect size benchmarks. Journal of Applied Psychology 2015;100:431.

67 Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat A-HS, Dellinger EP, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. N Engl J Med 2009;360:491–9.

68 Gaba DM, Howard SK, Flanagan B, Smith BE, Fish KJ, Botney R. 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. Acad Med 2014;89.

70 Gaba DM, Howard SK, Fish KJ, Smith BE, Sowb YA. Simulation-based training in anesthesia crisis resource management (ACRM): a decade of experience. Simulation & 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 Organiz Behav 1997:139–58.

72 LePine JA, Piccolo RF, Jackson CL, Mathieu JE, Saul JR. 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.

European Journal of Work and Organizational Psychology 2015;24:1–26.

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, J. B., & Manser, T. Training needs for adaptive coordination: Utilizing task analysis to identify coordination requirements in three different clinical settings. Gr & Org Manag 2018;43:504-527.

77 Eppich WJ, Mullan PC, Brett-Fleegler M, Cheng A. "Let's Talk About It": Translating Lessons From Health Care Simulation to Clinical Event Debriefings and Coaching Conversations. Clinical Pediatric Emergency Medicine 2016;17:200–11.

78 Schmutz JB, Eppich WJ. Promoting learning and patient care through shared reflection: A conceptual framework for team reflexivity in healthcare. Acad Med 2017;92:1555–63.

79 Salas E, DiazGranados D, Weaver SJ, King H. 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 El Bardissi 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.

89 Künzle B, Zala-Mezo E, Wacker J. Leadership in anaesthesia teams: the most effective leadership is shared. *Qual Saf Health Care* 2010;19:1–6.

90 Manojlovich M, DeCicco B. Healthy Work Environments, Nurse-Physician Communication, and Patients Outcomes. *Am J Crit Care* 2007;16:536–43.

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* 2007;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;32:41.

96 Siassakos D, Fox R, Crofts JF, et al. The management of a simulated emergency: Better teamwork, better performance. *Resuscitation*

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, et al. Leading to Recovery: Group Performance and Coordinative Activities in Medical Emergency Driven Groups. *Hum Perf* 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 H, Johnsen B, 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. 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.

*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, Schumacher, Legeret, et al.[80] | 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<br>Hands-on time within first 180 seconds |
| Brogaard, Kierkegaard, Hvidmand, et al.[60] | 2018 | To investigate the relationship between non-technical skills and clinical performance in obstetric teams | Video observation of obstetric teams (Obstetricians, obstetric 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, Kolbe, Wacker, et al.[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, Manser, Kolbe, et al.[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, Wacker, Grote, et al.[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, Min, Bridges[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, Giddings, Wilkinson, et al.[62] | 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, Mishra, Handa, et al.[63] | 2008 | To analyse the effects of surgical, aesthetic, 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 |
| Cooper, Wakelam[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 minutes | Survey about leadership behaviour using the Leadership Behaviour Description Questionnaire | Checklist based expert rating |
| Davenport, Henderson, Mosca, et al.[85] | 2007 | To measure the impact of organizational 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<br>Surgical mortality |
| El Bardissi, Wiegmann, Henrickson, 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, Chaboyer, Fairweather[87] | 2012 | To investigate how various human factors variables, extend the expected length of an operation | Live observation of surgical teams across 10 specialties | Structured observation of numbers of communication failures | Deviation from expected length of operation |

| | | | | | |
|---|---|---|---|---|---|
| Kolbe, Burtscher, Wacker, et al.[88] | 2012 | To test the relationship between speaking up and technical team performance in anaesthesia. | Observation of 2-person (nurse, resident) ad hoc anaesthesia teams performing simulated inductions of general anaesthesia with minor nonroutine events | Structured observation of speaking up behaviour | Checklist based expert rating |
| Kuenzle, Zala-Mezo, Wacker, et al.[89] | 2009 | To investigate shared leadership patterns during anaesthesia induction and to show how they are linked to team performance | Observation of 2-person (nurse, resident) ad hoc anaesthesia teams performing simulated inductions of general anaesthesia with a nonroutine event (asystole) | Structured observation of leadership behaviour | Reaction time to nonroutine event |
| Manojilovich, Antonakos, David, 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<br>Bloodstream infections<br>Pressure ulcers<br>Acute physiology and chronic health evaluation score |
| Manser, Bogdanovic, Clack, et al. [61] | 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, Müller, Marquardt, 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 |
| Mazzocco, Petitti, Fong, et al.[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, Catchpole, Dale, et al.[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, Hoffmann, Heimberg, et al.[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, Bristowe, Draycott, et al.[95] | 2012 | To investigate the relationship between patient satisfaction and communication | Video observation of teams (physicians, midwives) managing obstetric emergencies in secondary and tertiary maternity units | Structured observation of closed loop communication | Timely administration of magnesium sulphate |
| Siassakos, Fox, Crofts, et al.[96] | 2011 | To determine whether team performance in a simulated emergency is related to generic teamwork skills and behaviors | 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, Sexton, Lasky, et al.[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 | Compliance with Neonatal Resuscitation Program guidelines |
| Tschan, Semmer, Gautschi, et al.[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, Semmer, Gurtner, et al.[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 |
| Westli, Johnsen, Eid, et al.[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, El Bardissi, Dearani, et al.[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, Lasky, Dannemiller, et al.[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, Phillips-Bute, Petrusa, et al.[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, Fuerch, Halamek[104] | 2016 | To investigate the effect of standardized 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<br>Time to initiate positive pressure ventilation<br>Time to chest compression |

*Table 2.* Studies, effect sizes and moderator variables included in the meta-analytic database

| Authors | Year | Study goal | Setting | No. of teams | Professional composition | Team famil-iarity | Average team size | Task type | Patient realism | Perfor-mance measure |
|---|---|---|---|---|---|---|---|---|---|---|
| Amacher, Schumacher, Legeret, et al.[80] | 2017 | .11 | Emergency medicine | 72 | Uniprofessional | Experi-mental | 3 | Non-routine | Simulated | Process |
| Brogaard, Kierkegaard, Hvidmand, et al.[60] | 2018 | .43 | Obstetrics | 99 | Interprofessional | Real | 5 | Non-routine | Real | Process |
| Burtscher, Kolbe, Wacker, et al.[81] | 2011 | -.27 | Anaesthesia | 31 | Interprofessional | Experi-mental | 2 | Routine | Simulated | Process |
| Burtscher, Manser, Kolbe, et al.[82] | 2011 | .19 | Anaesthesia | 15 | Interprofessional | Experi-mental | 2 | Routine & non-routine | Simulated | Process |
| Burtscher, Wacker, Grote, et al.[83] | 2010 | .07 | Anaesthesia | 22 | Interprofessional | Real | 3 | Non-routine | Real | Process |
| Carlson, Min, Bridges[9][b] | 2009 | .83 | Emergency medicine | 44 | Uniprofessional | Experi-mental | 2.6 | Non-routine | Simulated | Process |
| Catchpole, Giddings, Wilkinson, et al.[62] | 2007 | .45† | Surgery | 24 | Interprofessional | Real | 9 | Non-routine | Real | Process |
| | 2007 | .29† | Surgery | 18 | Interprofessional | Real | 5 | Routine | Real | Process |
| Catchpole, Mishra, Handa, et al.[63] | 2008 | .36† | Surgery | 26 | Interprofessional | Real | | Routine | Real | Process |
| | 2008 | .09† | Surgery | 22 | Interprofessional | Real | | Routine | Real | Process |
| Cooper, Wakelam[84] | 1999 | .50 | General care | 20 | Interprofessional | Real | 4 | Routine | Real | Process |
| Davenport, Henderson, Mosca, et al.[85] | 2007 | .17 | Surgery | 52 | Interprofessional | Real | | Routine | Real | Outcome |
| El Bardissi, Wiegmann, Henrickson, et al.[86] | 2008 | .67 | Surgery | 31 | Interprofessional | Real | 7 | Routine | Real | Process |
| Gillespie, Chaboyer, Fairweather[87] | 2012 | .23 | Surgery | 160 | Interprofessional | Real | 6 | Routine | Real | Process |
| Kolbe, Burtscher, Wacker, et al.[88] | 2012 | .33 | Anaesthesia | 31 | Interprofessional | Real | 2 | Non-routine | Simulated | Process |
| Kuenzle, Zala-Mezo, Wacker, et al.[89] | 2009 | .56 | Anaesthesia | 12 | Interprofessional | Real | 2 | Routine | Simulated | Process |
| Manojilovich, Antonakos, David, et al.[90] | 2009 | .11 | Intensive care | 25 | Uniprofessional | Real | 36 | Routine | Real | Outcome |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Manser, Bogdanovic, Clack, et al.[61] | 2015 | .39 | Surgery | 19 | Interprofessional | Experi-mental | 5 | Routine | Simulated | Process |
| Marsch, Müller, Marquardt, et al.[91] | 2004 | .23 | Intensive care | 16 | Interprofessional | Experi-mental | 3 | Non-routine | Simulated | Process |
| Mazzocco, Petitti, Fong, et al.[92] | 2009 | .11 | Surgery | 293 | Interprofessional | Real | 6 | Routine | Real | Outcome |
| Mishra, Catchpole, Dale, et al.[93] | 2008 | .05 | Surgery | 26 | Interprofessional | Real | 6 | Routine | Real | Process |
| Schmutz, Hoffmann, Heimberg, et al.[94] | 2015 | .12 | Emergency medicine | 68 | Interprofessional | Real | 6 | Non-routine | Simulated | Process |
| Siassakos, Bristowe, Draycott, et al.[95] | 2012 | .66 | Obstetrics | 19 | Interprofessional | Real | 6 | Non-routine | Simulated | Process |
| Siassakos, Fox, Crofts, et al.[96] | 2011 | .55 | Emergency medicine/ obstetrics | 24 | Interprofessional | Experi-mental | 6 | Non-routine | Simulated | Process |
| Thomas, Sexton, Lasky, et al.[97] | 2006 | .23 | Neonatal care | 132 | Interprofessional | Real | 5 | Non-routine | Real | Process |
| Tschan, Semmer, Gautschi, et al.[98] | 2006 | .23 | Emergency medicine | 21 | Interprofessional | Experi-mental | 5 | Non-routine | Simulated | Process |
| Tschan, Semmer, Gurtner, et al.[99] | 2009 | .37 | Emergency medicine | 20 | Uniprofessional | Experi-mental | 2.65 | Non-routine | Simulated | Outcome |
| Westli, Johnsen, Eid, et al.[100] | 2010 | .18 | Emergency medicine | 27 | Interprofessional | Real | 5.1 | Non-routine | Simulated | Process |
| Wiegmann, El Bardissi, Dearani, et al.[101] | 2007 | .56 | Surgery | 31 | Interprofessional | Real | | Routine | Real | Process |
| Williams, Lasky, Dannemiller, et al.[102] | 2010 | .18 | Neonatal care | 12 | Interprofessional | Real | 5 | Non-routine | Real | Process |
| Wright, Phillips-Bute, Petrusa, et al.[103] | 2009 | .81 | General care | 9 | Uniprofessional | Experi-mental | 4 | Non-routine | Simulated | Process |
| Yamada, Fuerch, Halamek[104] | 2016 | .11 | Emergency medicine | 13 | Interprofessional | Experi-mental | 3 | Non-routine | Simulated | Process |

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

[b] Carlson, Min & Bridges has been identified as an outlier and therefore excluded from the analysis

*Table 3.* Meta-Analytic relationships between teamwork and clinical performance

| | N | k | r | 95% CI | 80% CR | Q | $I^2$ |
|---|---|---|---|---|---|---|---|
| Overall relationship | 1,390 | 31 | .28* | [.20 ; .35] | [.09 ; .45] | 53.7* | 46.0 |
| *Team characteristics* | | | | | | | |
| Professional composition | | | | | | | |
| Interprofessional | 1,264 | 27 | .28* | [.20 ; .36] | [.09 ; .46] | 47.1* | 48.2 |
| Uniprofessional | 126 | 4 | .28 | [-.01 ; .52] | [-.04 ; .54] | 6.5 | 47.1 |
| Team familiarity | | | | | | | |
| Experimental team | 240 | 10 | .25* | [.05 ; .43] | [-.05 ; .51] | 17.2* | 47.2 |
| Real team | 1,150 | 21 | .29* | [.20 ; .37] | [.12 ; .45] | 36.2* | 45.7 |
| Team size[a] | | | | | | | |
| *Task characteristics* | | | | | | | |
| Task type | | | | | | | |
| Routine task | 766 | 14 | .27* | [.12 ; .40] | [-.01 ; .50] | 30.9* | 65.0 |
| Non-routine task | 609 | 16 | .29* | [.20 ; .39] | [.16 ; .42] | 20.5 | 24.6 |
| *Methodological factors* | | | | | | | |
| Patient realism | | | | | | | |
| Real patient | 993 | 16 | .28* | [.18 ; .38] | [.10 ; .45] | 28.7* | 49.3 |
| Simulated patient | 397 | 15 | .28* | [.13 ; .41] | [.02 ; .50] | 25.0* | 44.6 |
| Performance measures | | | | | | | |
| Outcome performance | 390 | 4 | .13* | [.03 ; .23] | [.06 ; .19] | 1.3 | 0.0 |
| Process performance | 1,000 | 27 | .30* | [.21 ; .39] | [.10 ; .49] | 45.6* | 45.6 |

*Note*. $k$ = number of studies; $N$ = cumulative sample size (number of teams); $r$ = sample-size weighted correlation; CI = confidence interval; CR = credibility interval; $Q$ = test statistic for residual heterogeneity of the models; $I^2$ = % of total variability in the effect size estimates due to heterogeneity among true effects (vs. sampling error)

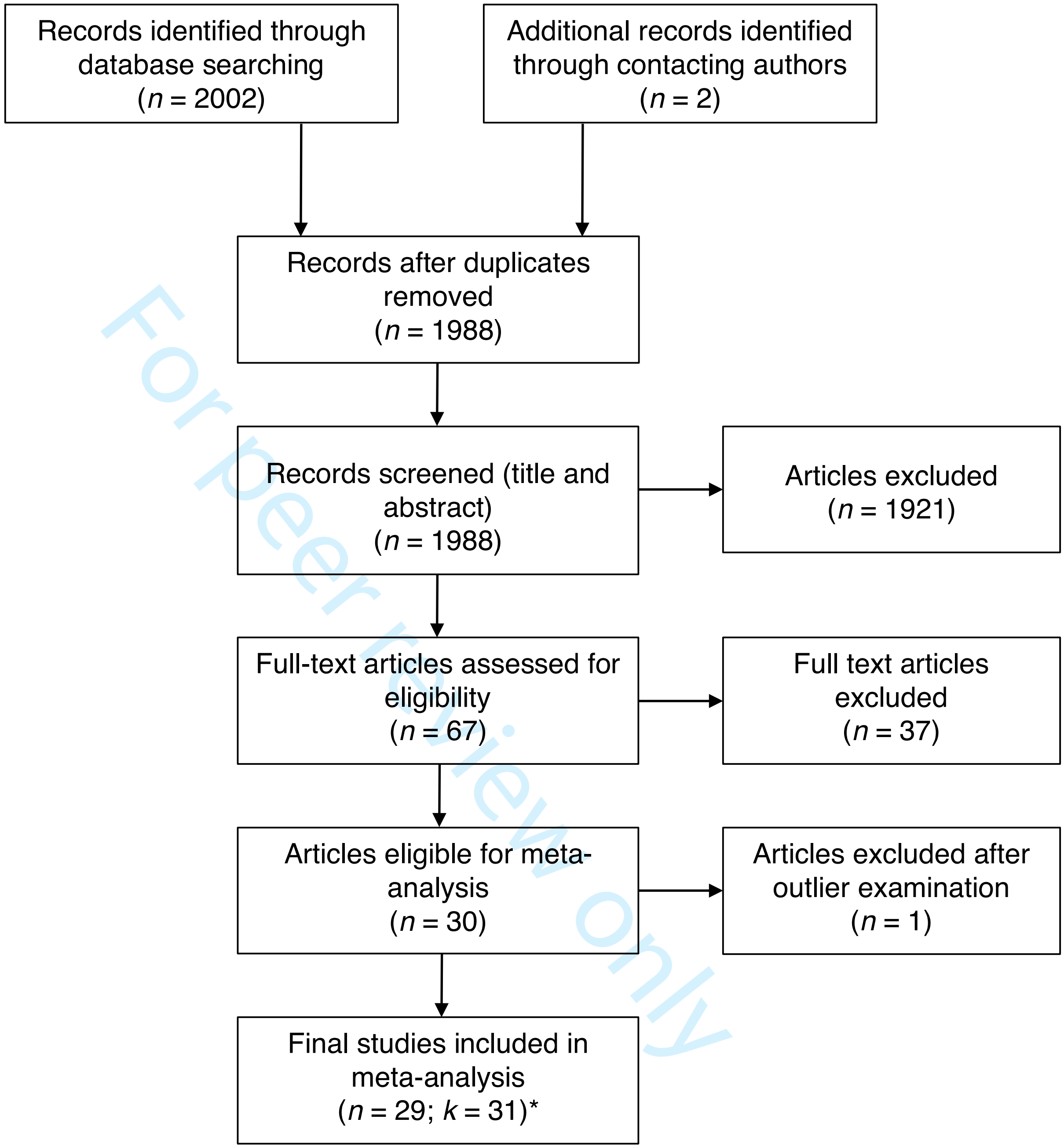
[a] Team size was entered as a continuous variable, therefore, no subgroup analyses exist

* $p < .05$.

## LEGENDS TO FIGURES

**Figure 1** Systematic literature search

**Figure 2** Relationship between teamwork processes and performance


Records identified through database searching
(n = 2002)
Additional records identified through contacting authors
(n = 2)
Records after duplicates removed
(n = 1988)
Records screened (title and abstract)
(n = 1988)
Articles excluded
(n = 1921)
Full-text articles assessed for eligibility
(n = 67)
Full text articles excluded
(n = 37)
Articles eligible for meta-analysis
(n = 30)
Articles excluded after outlier examination
(n = 1)
Final studies included in meta-analysis
(n = 29; k = 31)*

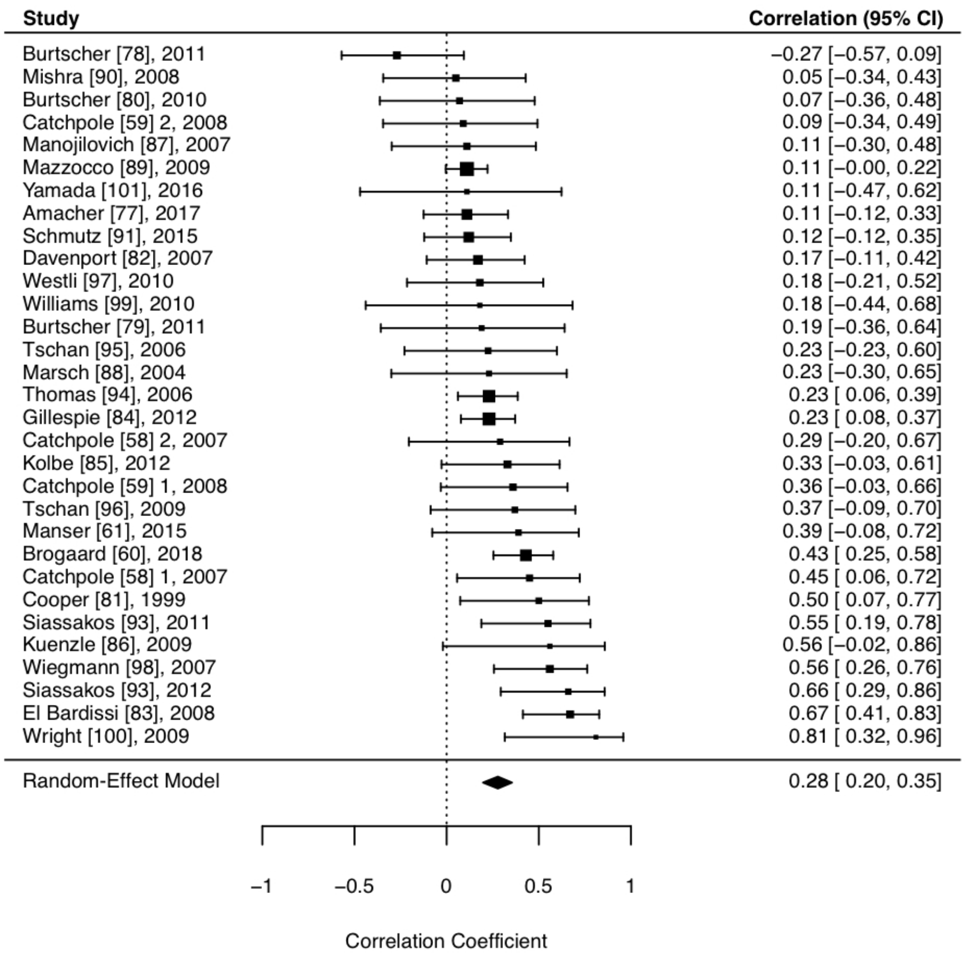



Figure 2. Relationship between teamwork processes and performance.

**SUPPLEMENTARY FILE**

Article: **How effective is teamwork really? The relationship between teamwork and performance in healthcare teams: A systematic review and meta-analysis**

Jan B. Schmutz, PhD, Laurenz L. Meier, PhD, Tanja Manser, PhD

**Search terms used for PubMed search**

(Teamwork[All Fields]) OR (team[All Fields] AND coordination[All Fields]) OR (team[All Fields] AND "decision making"[All Fields]) OR ((team[All Fields]) AND (communication[MeSH Terms]) OR (communication[All Fields])) OR (team[All Fields] AND leadership[All Fields]) AND (patient safety[MeSH Terms])

# PRISMA 2009 Checklist

| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| **TITLE** | | | |
| Title | 1 | Identify the report as a systematic review, meta-analysis, or both. | 1 |
| **ABSTRACT** | | | |
| Structured summary | 2 | Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number. | 2 |
| **INTRODUCTION** | | | |
| Rationale | 3 | Describe the rationale for the review in the context of what is already known. | 4-8 |
| Objectives | 4 | Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS). | 4-5 |
| **METHODS** | | | |
| Protocol and registration | 5 | Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number. | n/a |
| Eligibility criteria | 6 | Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. | 9-10 |
| Information sources | 7 | Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched. | 9 |
| Search | 8 | Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated. | 9 and supplemental material |
| Study selection | 9 | State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). | 9-10 |
| Data collection process | 10 | Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators. | 10 |
| Data items | 11 | List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made. | 9-12 |
| Risk of bias in individual studies | 12 | Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis. | 20 |
| Summary measures | 13 | State the principal summary measures (e.g., risk ratio, difference in means). | 13-19 / 23 |

# PRISMA 2009 Checklist

| | | | |
|---|---|---|---|
| Synthesis of results | 14 | Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., $I^2$) for each meta-analysis. | 20-23 |



| Section/topic | # | Checklist item | Reported on page # |
|---|---|---|---|
| Risk of bias across studies | 15 | Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies). | 20,26 |
| Additional analyses | 16 | Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified. | 22-23 |
| **RESULTS** | | | |
| Study selection | 17 | Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram. | 21 |
| Study characteristics | 18 | For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations. | 18-19 |
| Risk of bias within studies | 19 | Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12). | 18-19, 20 |
| Results of individual studies | 20 | For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot. | 13-19 |
| Synthesis of results | 21 | Present results of each meta-analysis done, including confidence intervals and measures of consistency. | 23 |
| Risk of bias across studies | 22 | Present results of any assessment of risk of bias across studies (see Item 15). | 23 |
| Additional analysis | 23 | Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]). | N/A |
| **DISCUSSION** | | | |
| Summary of evidence | 24 | Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers). | 24-27 |
| Limitations | 25 | Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias). | 26-27 |
| Conclusions | 26 | Provide a general interpretation of the results in the context of other evidence, and implications for future research. | 27 |
| **FUNDING** | | | |
| Funding | 27 | Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. | 28 |

# PRISMA 2009 Checklist

*From:* Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.