Associations of workflow disruptions in the operating room with surgical outcomes: a systematic review and narrative synthesis

Amelie Koch [1], Jacob Burns [2,3], Ken Catchpole [4], Matthias Weigl [1,5]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bmjqs-2019-010639).

For numbered affiliations see end of article.

Correspondence to
Amelie Koch, Institute and Outpatient Clinic for Occupational, Social and Environmental Medicine, Ludwig Maximilians University Munich Faculty of Medicine, Munich, Germany;
amelie.koch@med.uni-muenchen.de

Received 15 November 2019
Revised 27 April 2020
Accepted 28 April 2020
Published Online First 23 May 2020

To cite: Koch A, Burns J, Catchpole K, *et al*. *BMJ Qual Saf* 2020;**29**:1033–1045.

ABSTRACT

Background Performance in the operating room is an important determinant of surgical safety. Flow disruptions (FDs) represent system-related performance problems that affect the efficiency of the surgical team and have been associated with a risk to patient safety. Despite the growing evidence base on FDs, a systematic synthesis has not yet been published.

Objective Our aim was to identify, evaluate and summarise the evidence on relationships between intraoperative FD events and provider, surgical process and patient outcomes.

Methods We systematically searched databases MEDLINE, Embase and PsycINFO (last update: September 2019). Two reviewers independently screened the resulting studies at the title/abstract and full text stage in duplicate, and all inconsistencies were resolved through discussion. We assessed the risk of bias of included studies using established and validated tools. We summarised effects from included studies through a narrative synthesis, stratified based on predefined surgical outcome categories, including surgical process, provider and patient outcomes.

Results We screened a total of 20 481 studies. 38 studies were found to be eligible. Included studies were highly heterogeneous in terms of methodology, medical specialty and context. Across studies, 20.5% of operating time was attributed to FDs. Various other process, patient and provider outcomes were reported. Most studies reported negative or non-significant associations of FDs with surgical outcomes.

Conclusion Apart from the identified relationship of FDs with procedure duration, the evidence base concerning the impact of FDs on provider, surgical process and patient outcomes is limited and heterogeneous. We further provide recommendations concerning use of methods, relevant outcomes and avenues for future research on associated effects of FDs in surgery.

INTRODUCTION

Technical and organisational innovations have turned operating rooms (ORs) into highly complex and challenging working environments.[1] OR teams face multiple tasks simultaneously with high cognitive and technical demands.[2] Due to these and other challenges, the complex structure of OR work can lead to medical errors and suboptimal care.[3] Even minor decrements in OR team performance or small changes in the intraoperative environment can have consequences for surgical safety and patient outcomes.[4] OR system design efforts seek to identify preventable surgical errors and to improve intraoperative safety.[5]

Surgical 'flow disruptions' (FDs) represent a key challenge to OR team performance. Such events have been described as 'deviations of the natural progression of the operative procedure'.[6 7] They represent mismatches between the work demands and the configuration of the system to support the work, and range from interruptions (such as phone or beeper calls), to unexpected patient conditions or malfunctions of technical equipment. These small events potentially divert members of the surgical team from their primary task, increase cognitive workload and, as a result, may create a more serious, potentially harmful, situation.[1 8 9] Other industrial sectors, such as driving and aviation, have successfully addressed the challenges of distractions and task interferences, through conducting in-depth investigations and implementing tailored solutions.[4 10]

For the development of effective mitigation strategies and OR system interventions, a systematic assessment and thorough evaluation of the sequelae of FDs in ORs is necessary. The actual effects of FDs on the surgery itself, on providers and patient care are manifold.[11] Previous investigations found that FDs lead to higher stress levels, fatigue[12] and expand

procedure time, for example, in cardiac surgery, about 7% of operating time was associated with FD events.[13] Other studies reported larger delays up to 20%.[14 15] This considerable prolongation of the operating time results in substantially increased costs.[16] As current research suggests that different types of FDs may have specific effects—some are harmful, others beneficial and some essential for a seamless surgical workflow—detailed classification and systematic aggregation is essential.[11]

Despite the growing interest and literature on surgical FDs, there is, to the best of our knowledge, no systematic review that synthesises the current research base and considers the diversity of potential effects. In contrast to previous reviews,[6 11] our review further sought to encompass the broad spectrum of potential consequences as well as to systematically appraise the methodological quality of the study base. Establishing strong evidence provides a base for improving working conditions for the entire surgical team, optimising surgical safety and reducing intraoperative costs.[6] Identification of sociotechnical factors that potentially moderate the impact of FD events should contribute to safer surgical performance. Since the goal is not to eliminate all disruptions,[9] we need to know specific consequences and influencing factors of individual FD events in order to implement specific intraoperative interventions and effective surgical training.

Our systematic review aims to provide a comprehensive synthesis of the evidence on the association between intraoperative FDs and surgical outcomes. Apart from synthesising the current evidence base, we also seek to identify gaps in the current literature, from both a surgical and a methodological perspective.

METHODS

We conducted the systematic review in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses[17] and MOOSE Guidelines for Meta-Analyses and Systematic Reviews of Observational Studies.[18] We also consulted the Reporting Guideline for Synthesis without Meta-Analysis to ensure quality and transparency in reporting.[19] We developed a review protocol a priori, which included a detailed plan of steps, definitions of relevant constructs, eligibility criteria and data extraction procedures. This protocol was registered on 9 January 2019 (International prospective register of systematic reviews, PROSPERO, Nr. CRD42019120968).

Definition of FDs

We define surgical FDs as 'unexpected events causing a break in the natural progression of the operation'.[7 20] Visitors, unavailable or broken surgical equipment or training activities could be examples of typical FDs.[21] These include both temporary withdrawals of provider attention away from a primary (eg, dissection) to a secondary task (eg, answering a phone call), as well as actual breaks in task activity.[22] We excluded all scheduled activities such as consulting or completing checklists as well as preplanned surgical interventions. Continuous distractions or persistent stressors such as background noise, time pressure or fatigue were excluded, as they do not necessarily cause an immediate break in workflow.

Definition of 'surgical outcomes'

To capture all potentially relevant consequences of FDs, we included three different categories of intraoperative or postoperative outcomes: (1) surgical process outcomes (eg, duration of surgery and surgical performance), (2) team and provider outcomes (eg, workload and stress) and (3) patient outcomes (eg, infection rates, mortality and readmissions).

Search strategy and eligibility criteria

We searched MEDLINE (via PubMed), Embase, Google Scholar and PsycINFO (last updated search: 25 September 2019). Our search strategy included three key concepts: (A) 'operating room' (ie, setting), (B) 'flow disruptions' (ie, independent variable) and (C) 'surgical outcome' (ie, dependent variable). For each of these concepts, we created an individual search block by combining all synonyms using the Boolean operator 'OR'. The individual blocks were then combined using 'AND' (see online supplemental digital content 1). Additionally, we checked the references of all included studies for further relevant studies.

We applied the following criteria for the inclusion/exclusion of studies in our systematic review:

1. Setting: all studies must have assessed the main intraoperative phase of surgeries inside the OR. Investigations of surgical procedures outside ORs (eg, ambulatory interventions) or non-surgical procedures inside ORs were excluded.
2. Constructs of interest: a quantitative association between surgical FDs and surgical outcomes (see above) must have been measured and reported.
3. Original studies: only studies providing original, empirical and metric data were considered. Narrative literature research and systematic reviews were excluded.
4. Study design and methods: eligible were experimental studies (eg, randomised controlled trials (RCTs), cluster RCTs and non-RCTs, quasiexperimental studies (eg, controlled before-after studies and interrupted time series studies) and observational studies (eg, cohort studies, case–control studies and cross-sectional studies). (Expert) opinions, statements, case reports and case series were excluded; we did not apply any further restrictions concerning assessment methods for identification of FDs.
5. Availability: articles must be available in either English or German and must have been published in peer-reviewed journals between 1 January 2000 and 25

September 2019. We considered this time period since we aimed to capture the evidence for the 'modern' OR environment and account for the significant technological and equipment changes that occurred in the past two decades.

6. Where duplicate reports of the same study or cohort where identified, the most up-to-date or information-rich version was included.

Study selection

One author (AK) initially removed all duplicate reports. Subsequently, two assessors (AK and MW) independently screened all titles and abstracts. Subsequently, for those titles and abstracts deemed potentially relevant, these two assessors screened the full texts. Where full texts were not available online, authors were contacted. Reasons for exclusion were documented for each study at this stage of the selection process. Finally, all articles considered eligible were included. Any conflicts were resolved by discussion until consensus was reached. This also applies to inconsistencies during the following steps.

Data extraction

Two authors (MW and AK) independently extracted data for each included study by using a previously developed data sheet (see online supplemental digital content 2). The following data were extracted: authors, year of publication, title, country of origin, aim of the study, type of hospital/institution/department, setting, study design, sample characteristics, task or procedure characteristics, measures, FDs, surgical outcomes and size of quantitative associations between FDs and surgical outcomes. If data were not available in the publication, the respective field was marked with 'not reported'. Corresponding authors were contacted and asked for missing data to be sent. Although we did not receive additional data from all authors who were contacted, our calculations include some unpublished data that we were provided with at our request.

Risk of bias assessment

Two authors (MW and AK) independently assessed the risk of bias of included studies using established tools for risk of bias assessments: for RCTs, we applied the Cochrane Collaboration Risk of Bias (RoB 2) tool,[23] as recommended by the updated Cochrane Handbook for Systematic Reviews of Interventions.[24] In using this instrument, the assessor obtains a small, medium or high risk for bias for each assessed quality domain. Since all included RCTs used a cross-over design, we used the specific version for cross-over studies of the tool. For non-randomised studies (NRSs), we applied the Methodological Index for Non-Randomized Studies.[25] This tool enables a quality rating of included studies on a scale of 0–16 points (ie, high scores indicate low risk of bias).

Data synthesis

We planned a priori to conduct a meta-analysis where studies were sufficiently homogeneous to justify doing so. However, due to the substantial heterogeneity of included studies as well as missing data for certain outcomes, we were unable to do so. We summarised all outcomes through a narrative and graphical summary.[26] This summary was structured according to our predefined surgical outcome categories (ie, process, provider and patient outcomes). Our graphical synthesis involved the creation of two types of figures. First, for procedure duration, given that reported measures were fairly comparable across studies, we created an effect plot, summarising the proportion of the surgery procedure consisting of FDs. We considered calculating a weighted proportion for each study; however, given that SD was not reported in the majority of studies and that the range for the number of cases was relatively small, we calculated and reported an unweighted mean ('mean of means') across studies. Second, for all other outcomes, we created a bubble plot for summarising the investigated associations of FDs with provider, process and patient outcomes, arranging studies based on direction of effect. For a clearer overview and better structure, we clustered the outcomes of each study into meaningful outcome categories.

RESULTS

Eligible studies

A total of 20 481 articles were identified through database and hand searching. After duplicates were removed, we screened titles and abstracts of 13 355 articles. We subsequently reviewed the full texts of 114 articles. A further 76 articles were excluded (see figure 1). Finally, we included a total of 38 studies in our review.

Study characteristics

Included studies were highly heterogeneous in terms of both methodological approaches and study constructs. This arose through study design, measurement methods, setting, study participants and statistical analyses. In terms of investigated constructs, different definitions of FD variables and associated outcome constructs were applied (see table 1). As FDs encompass a potentially broad range of events that have been described in a number of ways, we gathered information on how the included studies defined and measured FD events (see table 1's column 'Operationalisation of FDs').

Included studies were published between 2007 and 2019 with more than the half of all studies published since 2016. Fifteen investigations were conducted in

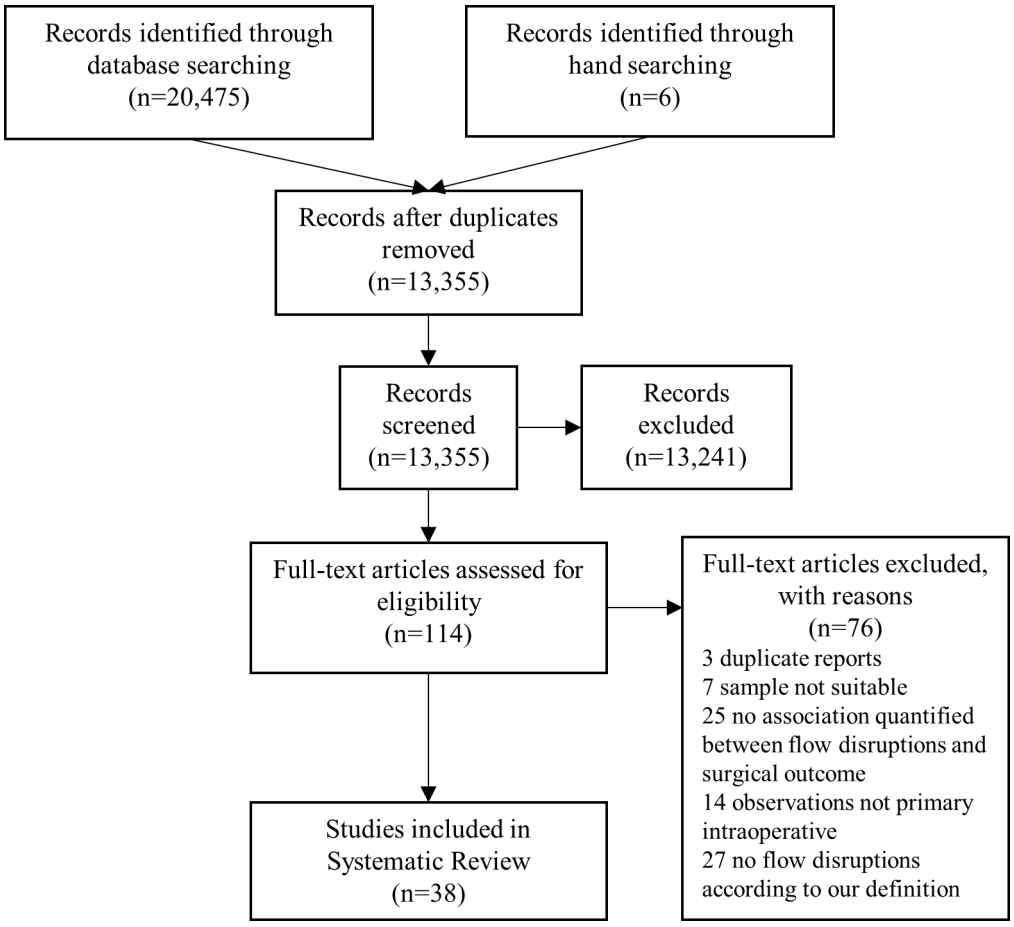


Figure 1 PRISMA flow chart of retrieved, screened and included studies. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Europe,[3 27–40] 15 in the USA and Canada,[7 13 14 21 22 41–50] 6 in Australia,[15 51–55] 1 in Asia[56] and 1 in New Zealand.[57] Studies were conducted in various specialties such as urology, gynaecology and cardiology. Overall, 1786 real surgical cases were observed in 31 observational studies, including open, minimally invasive and robotic-assisted procedures. Additionally, six simulation studies were conducted with 114 participants,[38 41–43 46 57] and one survey study with 194 participants.[31] Subjects of interest were primarily surgeons, nurses and anaesthetists, across different levels of experience. Among real-life observational studies, 12 obtained intraoperative data from the time patients were wheeled into the room until they were wheeled out.[13 21 28 32 33 36 40 47 51 55 56 58] Nine studies evaluated the period from incision to closure of the patient, that is, 'skin-to-skin'.[14 15 27 34 35 37 39 53 54] The remaining 10 studies selected other time periods or did not report the exact period.[3 7 22 29 30 44 48–50 52]

Risk of bias of included studies

Among the four included RCTs, one study was rated as having '*some concerns*'[42] and three were rated to be at '*high risk*'.[41 43 57] For the NRSs, quality ratings ranged from 5 points (highest risk of bias)[49 51] to 15 points (lowest risk of bias)[34] with an average rating of 9.29 points (see table 1).

Observed constructs of interest

Concerning FD events, almost all real-life observational studies considered nearly all types of FDs.[3 7 13–15 21 22 27–30 32–37 39 40 44 45 47 49–55] Thus, a wide range of FD events such as pager calls, equipment malfunctions, communication issues and door openings was covered. One study looked exclusively at disruptive cell phone calls.[56] Measurements of FDs differed substantially: some studies examined the frequency, some severity or interference and some the duration of FD events. Therefore, comparability of prevalence, rates and distributions across studies is limited. FD classification systems also differed significantly: an in-depth analysis of included classification systems and FD frequencies can be found elsewhere.[6] Furthermore, different methods were used to capture FDs: in 23 studies observers were present in the OR,[3 7 13 15 21 22 27–29 32–34 36 37 39 40 47 51 53–56 58] 7 studies used audio and video recordings[14 30 35 44 48–50] and 1 study combined both methods.[52] All simulation studies implemented more than one single FD event. These included distracting questions,[46 57] ringing cell phones,[37 41 57] pager interruptions,[42 43] side conversations,[41] sudden noise[41] or other minor events.

Table 1 Characteristics of included studies (first RCTs; second NRSs)

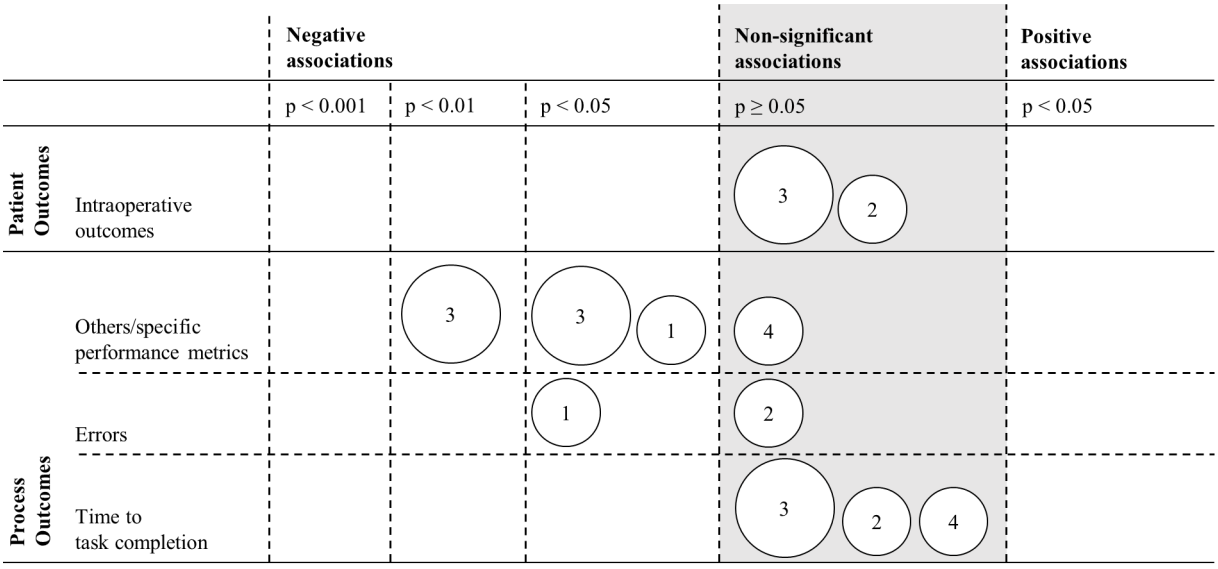
Included study	Quality rating	Study design	Simulated setting?	Providers/OR Team	Surgical task/ procedure	Sample size	Operationalisation of FDs	Surgical outcomes associated with FDs		
								Process	Provider	Patient
RCTs	**Cochrane Collaboration Risk of Bias Tool 2 (Quality Assessment Tool)**									
Feuerbacher 2012[41]	High risk.	Randomised, cross-over.	Yes.	Surgeons.	Laparoscopic cholecystectomy.	18*	All.	Errors and memory task.	–	–
Merry 2008[57]	High risk.	Randomised, cross-over.	Yes.	Anaesthetists.	Anaesthesia care.	10*	All.	Duration, errors and performance.	–	–
Murji 2016[42]	Some concerns.	Randomised, cross-over.	Yes.	Surgeons.	Laparoscopic salpingectomy.	30*	Pager interruptions.	Duration and performance.	–	–
Sujka 2018[43]	High risk.	Randomised, cross-over.	Yes.	Surgeons.	Laparoscopic cholecystectomy.	12*	Pager interruptions.	Duration and performance.	–	Complications.
NRSs	**Methodological Index of Non-Randomized Studies (Quality Assessment Tool)**									
Al-Hakim 2008[51]	5	Observational.	No.	OR team.	NR.	27†	All.	Duration.	–	–
Al-Hakim 2011[15]	6	Revelatory case study.	No.	OR team.	NR.	17†	All.	Duration.	–	–
Al-Hakim 2017[52]	7	Observational.	No.	OR team.	Laparoscopic urological procedures.	39†	All.	Duration.	–	–
Allers 2016[44]	7	Observational.	No.	OR team.	Robot-assisted radical prostatectomy.	10†	All.	Duration.	Workload.	–
Avidan 2017[56]	9	Observational.	No.	OR team.	NR.	52†	Phone calls.	Duration.	–	–
Blikkendaal 2018[27]	12	Observational.	No.	OR team.	Laparoscopic hysterectomy.	40†	All.	Duration.	–	Safety concerns.
Boquet 2017[22]	8	Observational.	No.	Anaesthesia team.	Anaesthesia care.	10†	All.	Duration.	–	–
Boquet 2017[13]	8	Observational.	No.	OR team.	Cardiovascular surgery.	24†	All.	Duration.	–	–
Campbell 2012[28]	8	Observational.	No.	Anaesthetists.	Anaesthesia care.	30†	All.	–	–	Patient consequences.
Catchpole 2016[21]	11	Observational.	No.	Surgeons.	Robotic surgery.	89†	All.	Duration.	–	–
Cohen 2016[45]	8	Observational.	No.	OR team.	Cardiac surgery.	15†	All.	Duration.	–	–
Cowan 2016[46]	8	Non-randomised, cross-over.	Yes.	Surgeons.	Diagnostic knee arthroscopy.	25*	Distracting questions.	Duration and performance.	–	–
Gillespie 2012[53]	9	Observational.	No.	OR team.	Various procedures.	160†	All.	Duration.	–	–
Gillespie 2012[54]	8	Observational.	No.	OR team.	Various procedures.	160†	All.	–	Miscommunication.	–
Gillespie 2017[55]	11	Observational.	No.	OR team.	Various procedures.	161†	All.	–	Non-technical skills.	–
Glarner 2017[14]	11	Observational.	No.	Surgeons.	Laparoscopic segmental colectomy.	10†	All.	Duration.	–	–
Healey 2007[29]	7	Observational.	No.	Urologists.	Urological surgery.	30†	All.	Duration.	–	–

Continued

Table 1 Continued

Included study	Quality rating	Study design	Simulated setting?	Providers/OR Team	Surgical task/ procedure	Sample size	Operationalisation of FDs	Surgical outcomes associated with FDs		
								Process	Provider	Patient
Henaux 2018[30]	10	Observational.	No.	Surgeons and nurses.	Anterior cervical discectomy and fusion.	12†	All.	Duration.	–	–
Jain 2016[47]	11	Observational.	No.	OR team.	Urological surgery.	32†	All.	Duration.	–	–
Jung 2019[48]	12	Cohort study.	No.	Surgeons.	Laparoscopic gastrointestinal surgery.	265†	All.	–	Distraction.	–
Pereira 2015[31]	11	Survey, cross-sectional.	No.	OR team.	Various procedures.	194*	All.	–	Attention failure and rumination.	–
Schraagen 2011[32]	11	Observational.	No.	OR team.	Paediatric cardiac surgery.	40†	All.	–	Teamwork.	Complications.
Sevdalis 2014[33]	10	Observational.	No.	OR team.	Urological surgery.	19†	All.	Performance.	–	–
Tiferes 2016[49]	5	Observational.	No.	OR team.	Urological surgery.	37†	All.	Duration.	–	–
Tschan 2015[34]	15	Observational.	No.	OR team.	Visceral surgery.	167†	All.	–	–	SSIs.
von Strauss 2016[35]	10	Observational.	No.	Surgeons and nurses.	Laparoscopic cholecystectomies.	20†	All.	Duration and costs.	–	–
Weber 2018[36]	9	Observational.	No.	OR team.	Robot-assisted radical prostatectomy.	40†	All.	Performance.	Workload, stress and distraction.	–
Weigl 2015[37]	10	Observational.	No.	Surgeons.	Various procedures.	56†	All.	None.	Workload.	–
Weigl 2016[38]	13	Non-randomised, cross-over.	Yes.	Surgeons.	Vertebroplasty.	19*	Phone call, patient discomfort.	Technical performance.	Workload.	–
Wheelock 2015[3]	9	Observational.	No.	OR team.	General surgery.	90†	All.	–	Stress, workload and teamwork.	–
Wiegmann 2007[7]	10	Observational.	No.	OR team.	Cardiac surgery.	31†	All.	Errors.	–	–
Willet 2018[39]	10	Observational.	No.	OR team.	Caesarean section.	56†	All.	Duration.	–	Patient safety.
Yoong 2015[40]	9	Observational.	No.	OR team.	Gynaecological surgery.	35†	All.	Duration.	–	Complications.
Zheng 2008[50]	8	Observational.	No.	OR team.	Laparoscopic antireflux surgery.	12†	All.	Duration.	–	–

Scale Range MINORS Quality Rating Tool 0–16 points: high scores indicate low risk of bias.
*Number of staff participants.
†Number of surgical cases.
FDs, flow disruptions; MINORS, Methodological Index for Non-Randomized Studies; NR, not reported; NRSs, non-randomised studies; OR, operating room; RCTs, randomised controlled trials; SSI, surgical site infections.



Note. Studies: 1=Feuerbacher 2012[41], 2=Merry 2008[57], 3=Murji 2016[42], 4=Sujka 2018[43]

Figure 2 Bubble plot for assessed associations with flow disruptions: study quality (larger bubble size indicates higher quality), p values, directions and types of associated outcomes (randomised controlled trials).

Concerning surgical outcomes, studies most frequently assessed process and provider outcomes, with only seven studies reporting patient outcomes.[27 28 32 34 39 40 43] Patient outcomes included intraoperative and postoperative complications, surgical-site infections (SSIs) or unspecified patient safety concerns. With regard to process outcomes, the duration of operating time attributable to FDs was most frequently recorded in 18 studies.[13–15 22 27 29 30 35 39 40 44 45 47 49–53] Further outcomes were errors, specific performance metrics (eg, instrument movement or bleedings) and costs. Surveyed provider outcomes were mental workload, teamwork, communication failures, non-technical skills, stress and perceived distraction.

Associations of FDs and surgical outcomes

Almost all studies identified either negative associations or no substantial associations of FDs with surgical outcomes. Two studies found a positive correlation, where more FDs were associated with a better outcome (ie, shorter duration of operating time and reduced workload).[3 21] In order to account for the inherent differences in risk of bias, we distinguish between RCTs (n=4) and NRSs (n=34) in the following.

Randomised controlled trials

Figure 2 shows that among the four included RCTs, the majority of investigated relationships revealed no significant association. These associations are graphically depicted using a bubble plot. The size of each bubble refers to the quality rating of each study (ie, larger bubble size indicates higher quality). The depicted associations refer to measures of correlations, group comparisons and regression analyses. Only one study showed a significantly increased task error rate when participants were disrupted compared with a control condition.[41] Others showed, for example, that participants made more unsafe decisions under the exposure to FDs.[42]

Non-randomised studies

With regard to the association between FDs and surgical process, 17 studies contributed data on the procedure duration, also often referred to as 'delays'. Figure 3 shows that the aggregated mean percentage of operating time attributed to FDs across these studies was 20.5%.[13–15 22 27 29 30 35 39 40 44 47 49–52 58] For most studies (n=11), no measure of variability was reported; a horizontal dotted line indicates this unreported variability.

Further assessed associations between FDs and surgical outcomes were highly diverse across included studies (see online supplemental digital content 3). Although associations were quite heterogeneous, overall mostly negative relationships were reported (see figure 4).

DISCUSSION

FDs are common in ORs, and they may have effects on surgical progress and patient safety.[59] We provide a systematic overview of the current state of empirical knowledge and report the first summary statistics for the prolongation of surgery time associated with FDs. Our review adds three specific contributions to the current evidence base:

First, we found one particularly important association of interest: our aggregated estimate suggests that on average about 20.5% of operating time is spent with FDs. To the best of our knowledge, our study is

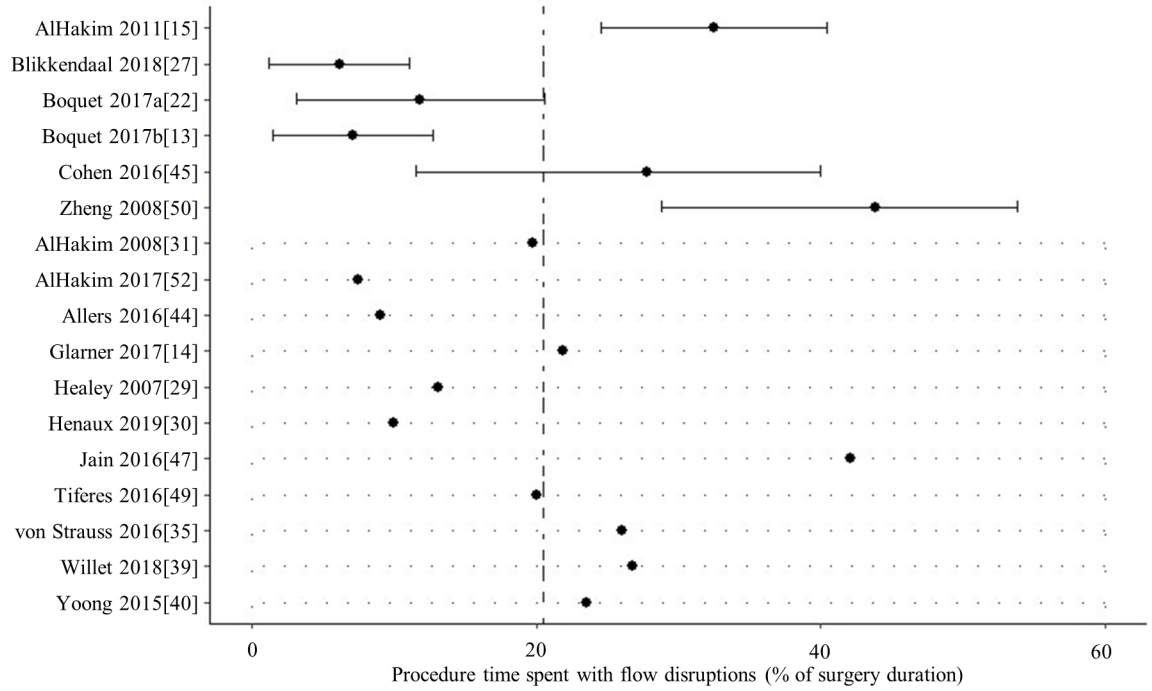
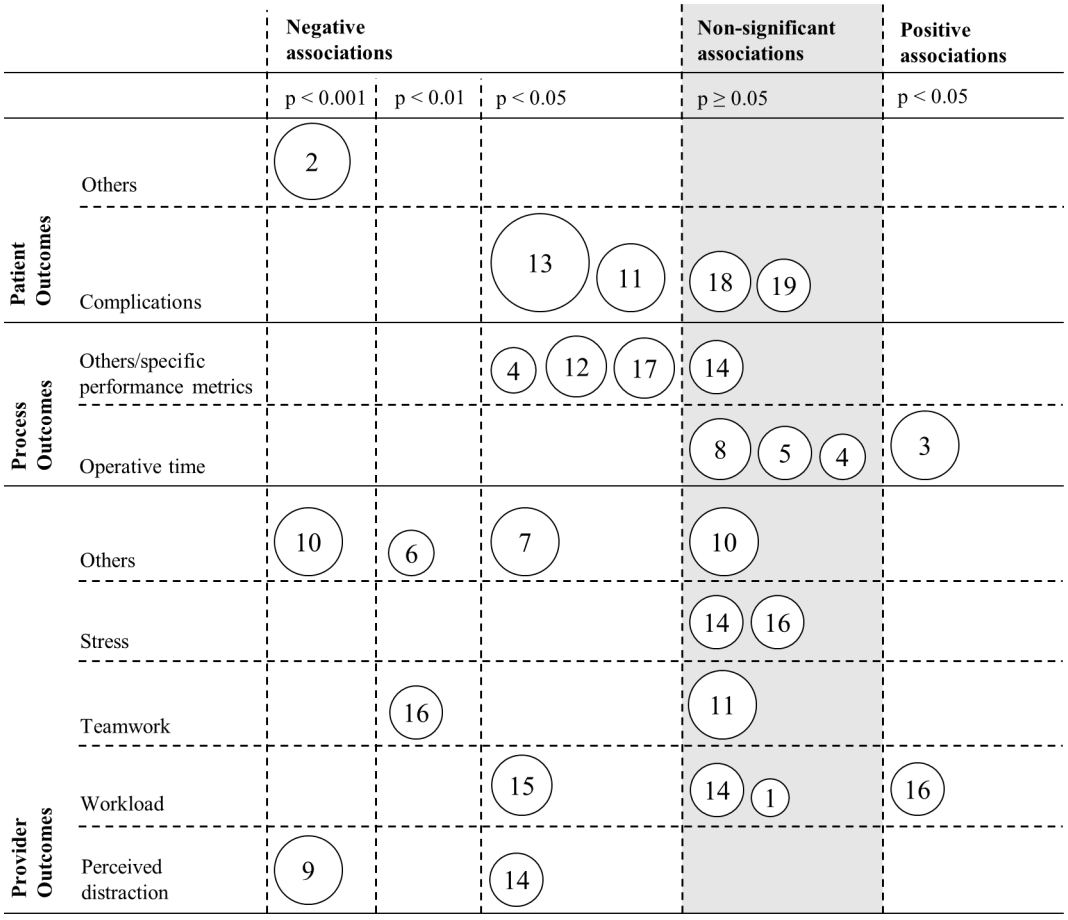


Figure 3 Mean percentages of procedure time attributed to flow disruption events, with SD. A horizontal dotted line indicates where measure of variability was not reported; the vertical dotted line represents the aggregated mean percentage.

the first to provide a summary statistic for this association, despite the diverse range of studies we reviewed. Even though we cannot determine the actual prolongation of surgeries caused by FDs, it seems that procedures with low levels of distractions tend to be shorter. Longer operating times mean longer working hours for the OR team, longer anaesthesia for the patient and higher costs for the hospital and should therefore be kept to a minimum.[16 60 61] It may appear logical that coping with frequent FDs prolongs procedures and causes delays; still, potential bidirectional influences should also be considered in the future. It is conceivable that procedure duration influences the occurrence of FDs, that is, likelihood that latent problems and inefficiencies occur with extended procedure time and accumulated delays.

We also observed that terminology of time-related outcomes was used differently across studies. Several studies used the term 'delay' to determine the procedure time spent with FDs. Others separated those FDs that actually cause a delay from those that do not, whereas others completely avoided the term 'delay'. Our findings suggest that the term 'delay' needs careful definition and operationalisation in observational studies. While some FDs may in fact demand that the primary surgeon steps away from the table with obviously observable time lags (eg, external phone calls), other frequent FDs such as 'door openings' may not cause delays at all. Moreover, even if an FD causes the procedure to pause or stop, this time lag can potentially be compensated for afterwards; for example, while the sterile nurse takes care of an equipment failure, the two surgeons use this pause as an opportunity to re-evaluate surgical goals, situation and progress. We would therefore recommend using the term 'delay' in this context carefully. We would instead suggest applying the term 'operating time spent with FDs'[13] and only using the term 'delay' if a time lag is directly and exclusively attributable to the occurring FDs.

Our second main finding was that the current evidence base on the sequelae of surgical FDs is heterogeneous, which limits inferences concerning associated risks. This refers to methods applied as well as to surveyed constructs and outcome variables. We originally aimed to conduct a meta-analysis on the impact of intraoperative FDs. We were able to categorise eligible studies into clusters of similar outcomes (eg, workload, teamwork, perceived distraction, errors and SSIs). However, incompatibility through conceptual and methodological heterogeneity precluded conducting a meta-analysis for all outcome categories. Beyond our attempts at aggregation, we noted that there is no common definition of the construct surgical 'flow disruptions'; varying terminology in identified studies include 'interruptions', 'disruptions', 'disturbances' and 'distractions'.[6] We applied a working definition of FDs, but it was not always possible to determine whether all events considered in a particular study adhered to our definition. We excluded studies of continuous distractions from this review yet acknowledge that permanent distractions (such as continuous noise or alarms) may impart disruptions and, subsequently, attentional failures or safety risks into the



Note. Studies: 1=Allers 2016[44], 2=Blikkendaal 2018[27], 3=Catchpole 2016[21], 4=Cowan 2016[46], 5=Gillespie 2012a[53], 5=Gillespie 2012b[54], 6=Gillespie 2017[55], 7=Henaux 2019[30], 8=Jung 2019[48], 9=Pereira 2015[31], 10=Schraagen 2011[32], 11=Sevdalis 2014[33], 12=Tschan 2015[34], 13=Weber 2018[36], 14=Weigl 2015[37], 15=Wheelock 2015[3], 16=Wiegmann 2007[7], 17=Willet 2018[39], 18=Yoong 2008[40].

Figure 4 Bubble plot for assessed associations with flow disruptions: study quality (larger bubble size indicates higher quality), p values, directions and types of associated outcomes (non-randomised studies).

OR. Moreover, FD events were operationalised and measured differently. Few studies merely recorded the incidence of FDs—some measure their duration, others evaluate their 'severity' or combine these characteristics. Furthermore, some authors define specific FD incidents that others have not considered.

Our third main finding was that the methodological quality of included studies was moderate to weak. Our quality assessment showed considerable deficiencies. For instance, in most studies, the sample size was not calculated in advance, and some studies are based on small sample sizes (n<15; both: surgical cases or staff participants). Of the RCTs, three-quarters were classified as potentially *high* risk of bias. We would therefore strongly recommend that future investigations adhere to guidelines and recommendations for high-quality research. Additionally, we identified considerably more observational field studies (n=31) than controlled trials, interventions and simulation studies (n=6). Naturalistic studies with cross-sectional designs that merely rely on observing and describing surgical FDs (ie, where trained observers are present in the OR and collect FD data or rely on video-based records) have the advantage to capture a wide range of intraoperative behaviours in situ and facilitate an understanding of the inherent and latent complexities of surgical work.[62] Nevertheless, randomised and controlled trials generally establish higher levels of internal validity than cross-sectional, descriptive studies, which are prone to several sources of bias.[63] Although descriptive studies inside the OR will be needed in the future to describe the actual state and 'work as done' in the OR, our results advocate for more high-quality studies, that is, interventions or naturalistic experiments on mitigating intraoperative FDs. Future attempts should also include study designs that allow for more robust inferences concerning the actual effects of FDs on provider, surgical as well as patient-reported outcomes. Simulations studies with manipulation of FD frequency and/or severity while controlling for surgeon's skill levels could be a good option for randomised trials in this field.

In addition to our main findings, there are two other insights. First, we found that mainly adverse

consequences of intraoperative FDs are reported. Unfortunately, it is generally unclear whether potential positive effects were not investigated or whether they were not reported. We identified two studies that reported potentially positive associations with surgical FDs.[3 21] Given this limited and inconsistent study base, it is premature to draw firm conclusions concerning potential beneficial effects of FDs in OR care systems. Further investigations addressing potentially favourable consequences of FDs would be highly desirable to prevent a unilateral state of research. For example, case-irrelevant humour during low complexity phases may leverage social tensions among an OR team. Second, as already stated elsewhere,[64] there is yet no clear indication that FDs have a harmful impact on patient outcomes or safety. However, since the evidence base is limited and only few postoperative patient outcomes have been investigated to date, we strongly recommend further investigations concerning actual patient outcomes. De Leval *et al*.[65] proposed that the surgical team might be resilient to the adverse effects and that therefore patient safety may not be compromised. In the light of our overall lack of substantive evidence of negative effects of FDs on surgical safety and patient care, we cannot draw firm conclusions on these outcomes besides case duration. Yet, the current state of literature does not refute the notion that some FDs may have a positive impact or that OR teams may be resilient to the negative impact of FDs. Future investigations in this field should consider the mechanisms of resiliency that OR teams use to prevent FDs from compromising patient outcomes.[48] This suggests that surgical processes in general, and FDs in particular, are likely to exhibit properties of complex adaptive systems, where process–outcome relationships may be indirect and mediated by a range of contextual factors.[66]

Limitations

Our first challenge was determining the inclusion of eligible studies based on our definition of FD events. Without a common consensus on what constitutes a FD, and with a wide range of alternative terms, it is difficult to identify all applicable investigations. Second, we did not search for studies published in languages other than English or German, and we did not search the grey literature. Third, we cannot rule out publication bias, and it is plausible that studies showing null effects or beneficial effects of FDs are under-reported. Furthermore, we found that some associations have been investigated but are not reported in the respective publications. We tried to obtain additional data from the authors, but these efforts were not always successful. Our findings need to be interpreted carefully since qualities and evidence levels of included studies differed substantially. We included both randomised and non-randomised studies, and shortcomings in methods and designs occurred across studies. Included studies consisted of small samples, some conducted retrospective data analyses and blinding (for observers as well as observed providers) was rarely applied.

Implications

This review provides essential implications for research as well as for surgical practice. First, many studies focused on specific outcomes without considering influencing factors. Individual relationships of FDs with outcomes are surveyed, yet the complexity and dynamics of surgical work is largely neglected. We recommend future investigations to apply a comprehensive view that considers provider and contextual factors; for example, a surgeon's skill level in coping with FDs, disruption handling strategies or organisational-level influences.[67] Second, it is evident that teamwork and team processes play a major role in surgical performance and may alter the impact of FDs.[68] 'Team familiarity' as well as team cognition factors may serve as protective factors.[44 55] For this reason, we recommend expanding focus from individual OR professionals to comprehensive assessments of the whole OR team. Third, we found that almost every research group used their own classification and evaluation systems to categorise intraoperative FDs. Future studies that scrutinise FDs and associated effects should strive for the use of validated observational tools in order to allow comparability across studies, procedures, settings and contexts.

Concerning implications for surgical practice, our results corroborate previous claims that most consequences of FDs inside the OR are negative in nature. Thus, efforts to mitigate FDs in the OR would be beneficial for provider cognition and safety. Organisational, teamwork, work or process redesign interventions that support a less disruptive work environment are highly recommended.[3 33 51] So far, the greatest impact of FDs in the OR was reported for the operating team and its individual members (eg, for workload and stress outcomes). Suitable interventions should therefore start here. One idea is to reduce the effects of FDs through targeted training of providers.[6] Kolodzey *et al*[5] recommended, in order to encourage team resilience, to focus interventions on the training of core competences such as teaching, effective teamwork and leadership. Complementary approaches should address organisational and external conditions to avoid unnecessary disruptions from outside the OR, such as case-irrelevant phone calls due to poor intradepartmental coordination and planning.

CONCLUSIONS

In the light of efforts to provide safe and efficient surgical care in the OR, we need an evidence-based understanding of the nature and potential effects of FDs that occur during the procedure. Our systematic review revealed that across all surgical disciplines a large proportion of operating time is associated with

distractions. Furthermore, we were able to demonstrate the extent to which FDs are associated with intraoperative and postoperative outcomes. Despite the wide range of available research, we identified major gaps that need to be addressed in future investigations of FDs. In order to strengthen the quality of studies on surgical FDs as well as to improve the comparability of their empirical findings, we recommend considering previously established definitions and methods. It appears that in many cases a reduction of FDs can be expected to have a positive effect on OR safety, but this hypothesis still requires some further elaboration and broader empirical confirmation.

Author affiliations
[1]Institute and Clinic for Occupational, Social and Environmental Medicine, LMU University Hospital, Munich, Germany
[2]Institute for Medical Information Processing, Biometry, and Epidemiology - IBE, Munich, Germany
[3]Pettenkofer School of Public Health, Munich, Germany
[4]Department of Anesthesia and Perioperative Medicine, Medical University of South Carolina (MUSC), Charleston, South Carolina, USA
[5]Division of Surgery and Cancer, Imperial College, London, UK

Twitter Ken Catchpole @KenCatchpole

Contributors AK and MW developed study objectives, design, conception and planning. AK and MW carried out studies search and selection, data extraction, analysis and interpretation. JB and KC contributed to data analyses, data presentation and interpretation. AK and MW were responsible for the first draft of the publication. All authors were involved equally in writing and shaping the final version of the manuscript.

Funding This study was funded by Munich Centre for Health Sciences (MC Health).

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. The datasets generated during and analysed during the current study are available from the corresponding author on reasonable request.

ORCID iDs
Amelie Koch http://orcid.org/0000-0002-4389-2868
Ken Catchpole http://orcid.org/0000-0003-4073-3025
Matthias Weigl http://orcid.org/0000-0003-2408-1725

REFERENCES

1 Palmer G, Abernathy JH, Swinton G, *et al*. Realizing improved patient care through human-centered operating room design: a human factors methodology for observing flow disruptions in the cardiothoracic operating room. *Anesthesiology* 2013;119:1066–77.
2 Ghazanfar MA, Cook M, Tang B, *et al*. The effect of divided attention on novices and experts in laparoscopic task performance. *Surg Endosc* 2015;29:614–9.
3 Wheelock A, Suliman A, Wharton R, *et al*. The impact of operating room distractions on stress, workload, and teamwork. *Ann Surg* 2015;261:1079–84.
4 Jung JJ, Jüni P, Lebovic G, *et al*. First-Year analysis of the operating room black box study. *Ann Surg* 2020;271:122–7.
5 Kolodzey L, Trbovich P, Kashfi A, *et al*. System factors affecting intraoperative risk and resilience: applying a novel integrated approach to study surgical performance and patient safety. *Ann Surg* 2019. doi:10.1097/SLA.0000000000003280
6 Mentis HM, Chellali A, Manser K, *et al*. A systematic review of the effect of distraction on surgeon performance: directions for operating room policy and surgical training. *Surg Endosc* 2016;30:1713–24.
7 Wiegmann DA, ElBardissi AW, Dearani JA, *et al*. Disruptions in surgical flow and their relationship to surgical errors: an exploratory investigation. *Surgery* 2007;142:658–65.
8 Antoniadis S, Passauer-Baierl S, Baschnegger H, *et al*. Identification and interference of intraoperative distractions and interruptions in operating rooms. *J Surg Res* 2014;188:21–9.
9 Shouhed D, Gewertz B, Wiegmann D, *et al*. Integrating human factors research and surgery: a review. *Arch Surg* 2012;147:1141–6.
10 McGowan G, Jawaheer L, Young D, *et al*. Quiet please! effect of distraction on simulated posterior segment surgical performance. *Graefes Arch Clin Exp Ophthalmol* 2018;256:519–23.
11 Bretonnier M, Michinov E, Morandi X, *et al*. Interruptions in surgery: a comprehensive review. *J Surg Res* 2020;247:190–6.
12 Al-Hakim L, Arora S, Sevdalis N. Impact of disruptions on anaesthetic workflow during anaesthesia induction and patient positioning: a prospective study. *Eur J Anaesthesiol* 2016;33:581–7.
13 Boquet AJ, Cohen TN, Reeves ST, *et al*. Flow disruptions impacting the surgeon during cardiac surgery: defining the boundaries of the error space. *Perioper Care Oper Room Manag* 2017;7:1–6.
14 Glarner CE, Law KE, Zelenski AB, *et al*. Resident training in a teaching hospital: how do attendings teach in the real operative environment? *Am J Surg* 2017;214:141–6.
15 Al-Hakim L. The impact of preventable disruption on the operative time for minimally invasive surgery. *Surg Endosc* 2011;25:3385–92.
16 von Strauss Und Torney M, Aghlmandi S, Zeindler J, *et al*. High-Resolution standardization reduces delay due to workflow disruptions in laparoscopic cholecystectomy. *Surg Endosc* 2018;32:4763–71.
17 Moher D, Shamseer L, Clarke M, *et al*. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
18 Stroup DF, Berlin JA, Morton SC, *et al*. Meta-Analysis of observational studies in epidemiology: a proposal for reporting. meta-analysis of observational studies in epidemiology (moose) group. *JAMA* 2000;283:2008–12.
19 Campbell M, McKenzie JE, Sowden A, *et al*. Synthesis without meta-analysis (swim) in systematic reviews: reporting guideline. *BMJ* 2020;368:l6890.
20 Shouhed D, Blocker R, Gangi A, *et al*. Flow disruptions during trauma care. *World J Surg* 2014;38:314–21.
21 Catchpole K, Perkins C, Bresee C, *et al*. Safety, efficiency and learning curves in robotic surgery: a human factors analysis. *Surg Endosc* 2016;30:3749–61.
22 Boquet A, Cohen T, Diljohn F, *et al*. A theoretical model of flow disruptions for the anesthesia team during cardiovascular surgery. *J Patient Saf* 2017. doi:10.1097/PTS.0000000000000406

23 Sterne JAC, Savović J, Page MJ, *et al*. Rob 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:l4898.
24 Munder T, Barth J. Cochrane's risk of bias tool in the context of psychotherapy outcome research. *Psychother Res* 2018;28:347–55.
25 Slim K, Nini E, Forestier D, *et al*. Methodological index for non-randomized studies (minors): development and validation of a new instrument. *ANZ J Surg* 2003;73:712–6.
26 Higgins JPT, Thomas J, Li T, *et al*. *Cochrane Handbook for systematic reviews of interventions version 6.0 (updated July 2019)*, 2019.
27 Blikkendaal MD, Driessen SRC, Rodrigues SP, *et al*. Measuring surgical safety during minimally invasive surgical procedures: a validation study. *Surg Endosc* 2018;32:3087–95.
28 Campbell G, Arfanis K, Smith AF. Distraction and interruption in anaesthetic practice. *Br J Anaesth* 2012;109:707–15.
29 Healey AN, Primus CP, Koutantji M. Quantifying distraction and interruption in urological surgery. *Qual Saf Health Care* 2007;16:135–9.
30 Henaux P-L, Michinov E, Rochat J, *et al*. Relationships between expertise, crew Familiarity and surgical workflow disruptions: an observational study. *World J Surg* 2019;43:431–8.
31 Pereira D, Müller P, Elfering A. Workflow interruptions, social stressors from supervisor(s) and attention failure in surgery personnel. *Ind Health* 2015;53:427–33.
32 Schraagen JM, Schouten T, Smit M, *et al*. A prospective study of paediatric cardiac surgical microsystems: assessing the relationships between non-routine events, teamwork and patient outcomes. *BMJ Qual Saf* 2011;20:599–603.
33 Sevdalis N, Undre S, McDermott J, *et al*. Impact of intraoperative distractions on patient safety: a prospective descriptive study using validated instruments. *World J Surg* 2014;38:751–8.
34 Tschan F, Seelandt JC, Keller S, *et al*. Impact of case-relevant and case-irrelevant communication within the surgical team on surgical-site infection. *Br J Surg* 2015;102:1718–25.
35 von Strauss Und Torney M, Dell-Kuster S, Hoffmann H, *et al*. Microcomplications in laparoscopic cholecystectomy: impact on duration of surgery and costs. *Surg Endosc* 2016;30:2512–22.
36 Weber J, Catchpole K, Becker AJ, *et al*. Effects of flow disruptions on mental workload and surgical performance in robotic-assisted surgery. *World J Surg* 2018;42:3599–607.
37 Weigl M, Antoniadis S, Chiapponi C, *et al*. The impact of intra-operative interruptions on surgeons' perceived workload: an observational study in elective general and orthopedic surgery. *Surg Endosc* 2015;29:145–53.
38 Weigl M, Stefan P, Abhari K, *et al*. Intra-Operative disruptions, surgeon's mental workload, and technical performance in a full-scale simulated procedure. *Surg Endosc* 2016;30:559–66.
39 Willett M, Gillman O, Shin E, *et al*. The impact of distractions and interruptions during cesarean sections: a prospective study in a London teaching hospital. *Arch Gynecol Obstet* 2018;298:313–8.
40 Yoong W, Khin A, Ramlal N, *et al*. Interruptions and distractions in the gynaecological operating theatre: irritating or dangerous? *Ergonomics* 2015;58:1314–9.
41 Feuerbacher RL, Funk KH, Spight DH, *et al*. Realistic distractions and interruptions that impair simulated surgical performance by novice surgeons. *Arch Surg* 2012;147:1026–30.
42 Murji A, Luketic L, Sobel ML, *et al*. Evaluating the effect of distractions in the operating room on clinical decision-making and patient safety. *Surg Endosc* 2016;30:4499–504.
43 Sujka JA, Safcsak K, Bhullar IS, *et al*. Simulation-Based testing of Pager interruptions during laparoscopic cholecystectomy. *J Surg Educ* 2018;75:1351–6.
44 Allers JC, Hussein AA, Ahmad N, *et al*. Evaluation and impact of workflow interruptions during robot-assisted surgery. *Urology* 2016;92:33–7.
45 Cohen TN, Cabrera JS, Sisk OD, *et al*. Identifying workflow disruptions in the cardiovascular operating room. *Anaesthesia* 2016;71:948–54.
46 Cowan JB, Seeley MA, Irwin TA, *et al*. Computer-Simulated arthroscopic knee surgery: effects of distraction on resident performance. *Orthopedics* 2016;39:e240–5.
47 Jain M, Fry BT, Hess LW, *et al*. Barriers to efficiency in robotic surgery: the resident effect. *J Surg Res* 2016;205:296–304.
48 Jung JJ, Elfassy J, Grantcharov T. Factors associated with surgeon's perception of distraction in the operating room. *Surg Endosc*;105.
49 Tiferes J, Hussein AA, Bisantz A, *et al*. The loud surgeon behind the Console: understanding team activities during robot-assisted surgery. *J Surg Educ* 2016;73:504–12.
50 Zheng B, Martinec DV, Cassera MA, *et al*. A quantitative study of disruption in the operating room during laparoscopic antireflux surgery. *Surg Endosc* 2008;22:2171–7.
51 Hakim LA. Surgical disruption: information quality perspective. *IJIQ* 2008;2:192.
52 Al-Hakim L, Xiao J, Sengupta S. Ergonomics perspective for identifying and reducing internal operative flow disruption for laparoscopic urological surgery. *Surg Endosc* 2017;31:5043–56.
53 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.
54 Gillespie BM, Chaboyer W, Fairweather N. Interruptions and miscommunications in surgery: an observational study. *Aorn J* 2012;95:576–90.
55 Gillespie BM, Harbeck E, Kang E, *et al*. Correlates of non-technical skills in surgery: a prospective study. *BMJ Open* 2017;7:e014480.
56 Avidan A, Yacobi G, Weissman C, *et al*. Cell phone calls in the operating Theater and staff Distractions: an observational study. *J Patient Saf* 2019;15:e52–5.
57 Merry AF, Weller JM, Robinson BJ, *et al*. A simulation design for research evaluating safety innovations in anaesthesia*. *Anaesthesia* 2008;63:1349–57.
58 Cohen TN, Cabrera JS, Sisk OD, *et al*. Identifying workflow disruptions in the cardiovascular operating room. *Anaesthesia* 2016;71:948–54.
59 Wiegmann DA, Sundt TM. Workflow disruptions and surgical performance: past, present and future. *BMJ Qual Saf* 2019;28:260–2.
60 Childers CP, Showen A, Nuckols T, *et al*. Interventions to reduce intraoperative costs: a systematic review. *Ann Surg* 2018;268:48–57.
61 Kobayashi K, Ando K, Ito K, *et al*. Factors associated with extension of the scheduled time for spine surgery. *Clin Neurol Neurosurg* 2018;169:128–32.
62 Catchpole K, Neyens DM, Abernathy J, *et al*. Framework for direct observation of performance and safety in healthcare. *BMJ Qual Saf* 2017;26:1015–21.

63 Mercuri M, Baigrie B, Upshur REG. Going from evidence to recommendations: can grade get us there? *J Eval Clin Pract* 2018;24:1232–9.
64 Sanderson P, McCurdie T, Grundgeiger T. Interruptions in health care: assessing their connection with error and patient harm. *Hum Factors* 2019;61:1025–36.
65 de Leval MR, Carthey J, Wright DJ, *et al*. Human factors and cardiac surgery: a multicenter study. *J Thorac Cardiovasc Surg* 2000;119:661–72.
66 Braithwaite J, Clay-Williams R, Nugus P, *et al*. Health care as a complex adaptive system. In: Hollnagel E, Braithwaite J, Wears RL, *et al*, eds. *Resilient health care*. United Kingdom: Ashgate Publishing, 2013: 57–73.
67 Blikkendaal MD, Driessen SRC, Rodrigues SP, *et al*. Surgical flow disturbances in dedicated minimally invasive surgery suites: an observational study to assess its supposed superiority over conventional suites. *Surg Endosc* 2017;31:288–98.
68 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.