

Impact of fatigue in surgeons on performance and patient outcome: systematic review

Iris M. Reijmerink^{1*} , Maarten J. van der Laan¹ , J. K. Götz Wietasch², Lotty Hooft³ and Fokie Cnossen⁴

¹Department of Surgery, University Medical Centre Groningen, University of Groningen, Groningen, The Netherlands

²Department of Anaesthesiology, University Medical Centre Groningen, University of Groningen, Groningen, The Netherlands

³Cochrane Netherlands, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

⁴Department of Artificial Intelligence, Bernoulli Institute of Mathematics, Computer Science and Artificial Intelligence, University of Groningen, Groningen, The Netherlands

*Correspondence to: Iris M. Reijmerink, Department of Surgery (Division of Vascular Surgery), University Medical Centre Groningen, University of Groningen, Hanzeplein 1, Groningen, 9713 GZ, The Netherlands (e-mail: i.m.reijmerink@umcg.nl)

Abstract

Background: While fatigue is an inevitable aspect of performing surgical procedures, lack of consensus remains on its effect on surgical performance. The aim of this systematic review was to assess the effect of non-muscular fatigue on surgical outcome.

Methods: MEDLINE and Embase were searched up to 17 January 2023. Studies on students, learning, duty-hour restrictions, muscle fatigue, non-surgical or subjective outcome, the weekend effect, or time of admission were excluded. Studies were categorized based on real-life or simulated surgery. The Cochrane risk-of-bias tool was used to assess RCTs and the Newcastle–Ottawa scale was used to assess cohort studies. Due to heterogeneity among studies, data pooling was not feasible and study findings were synthesized narratively.

Results: From the 7251 studies identified, 134 studies (including 1 684 073 cases) were selected for analysis (110 real-life studies and 24 simulator studies). Of the simulator studies, 46% (11 studies) reported a deterioration in surgical outcome when fatigue was present, using direct measures of fatigue. In contrast, only 35.5% (39 studies) of real-life studies showed a deterioration, observed in only 12.5% of all outcome measures, specifically involving aggregated surgical outcomes.

Conclusion: Almost half of simulator studies, along with one-third of real-life studies, consistently report negative effects of fatigue, highlighting a significant concern. The discrepancy between simulator/real-life studies may be explained by heightened motivation and effort investment in real-life studies. Currently, published fatigue and outcome measures, especially in real-life studies, are insufficient to fully define the impact of fatigue on surgical outcomes due to the absence of direct fatigue measures and crude, post-hoc outcome measures.

Introduction

Muscular and non-muscular fatigue are an inevitable aspect of performing surgical procedures¹, with complex procedures and excessive working hours contributing to fatigue^{1,2}. The negative effects of fatigue have even been compared to operating a vehicle with a blood alcohol concentration exceeding the legal limit³. While in other professions, for example in the aviation industry, fatigue has been incontrovertibly linked to diminished job functioning and an elevated incidence of errors, the effects of non-muscular fatigue, henceforth referred to simply as 'fatigue', on surgical performance are still a subject of controversy^{4–6}. Gawande *et al.*⁷ demonstrated that 33% of intraoperative errors were attributable to excessive fatigue, yet Ellman *et al.*⁸ failed to demonstrate a discernible impact of fatigue on surgical outcomes. McCormick *et al.*⁹ found that surgeons can be fatigued for approximately 50% of their waking hours, underscoring the importance of evaluating the potential impact of fatigue on performance to ensure high-quality surgical care. The aim of this systematic review was to assess

the impact of non-muscular surgeon fatigue on surgical performance, as measured through both real-life surgery and simulated surgery outcomes.

Methods

The present study was performed in accordance with PRISMA guidelines and was registered in PROSPERO (CRD42022290704).¹⁰

Literature search

MEDLINE and Embase were searched up to 17 January 2023. A search strategy was created by one reviewer (I.M.R.) in consultation with a clinical librarian (*Appendix S1*), using Medical Subject Headings terms and additional free entry terms for non-muscular fatigue, the surgical profession, and clinical competence, and then performed. Reference lists of selected articles and related reviews were screened for other relevant articles.

Received: August 31, 2023. Revised: October 29, 2023. Accepted: November 05, 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of BJS Society Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Inclusion and exclusion criteria

Inclusion and exclusion criteria were established by three study team members (I.M.R., M.J.v.d.L., and F.C.). One reviewer (I.M.R.) applied these criteria, with uncertainties resolved through discussion with the study team.

Included studies examined the impact of non-muscular fatigue in qualified surgeons and surgical residents in real-life surgery or simulated surgical sessions. In real-life studies, surgical outcome was assessed by patient-based outcome measures, such as number of complications. In simulator studies, surgical outcome was assessed by surgeon-based outcome measures, consisting of either cognitive function or psychomotor function outcome measures, assessed by means of simulated tasks.

Studies on students, learning, duty-hour restrictions, muscle fatigue, non-surgical or subjective outcome, the weekend effect, or time of admission were excluded, as were studies with non-English full text or article types other than original articles. Time-related outcome measures were not included in the analysis, as a decrease

in time to task completeness could indicate a learning effect, as well as decreased attention to detail^{11,12}.

Data extraction

Data extraction was performed by one reviewer (I.M.R.) and validated by a second reviewer (Neda Ansari (see the Acknowledgements section)) in a random 10% sample of extracted data. Results from both investigators were compared and any discrepancies were resolved. As no major discrepancies were found, no further verification was undertaken.

Extracted data included research design, type and number of participants, surgical skill studied, setting (that is real-life or simulator), surgical outcome measures studied, fatigue indicators used, and main study findings.

Quality assessment

One reviewer (I.M.R.) assessed the quality of each included study. The Cochrane risk-of-bias tool was used to assess the

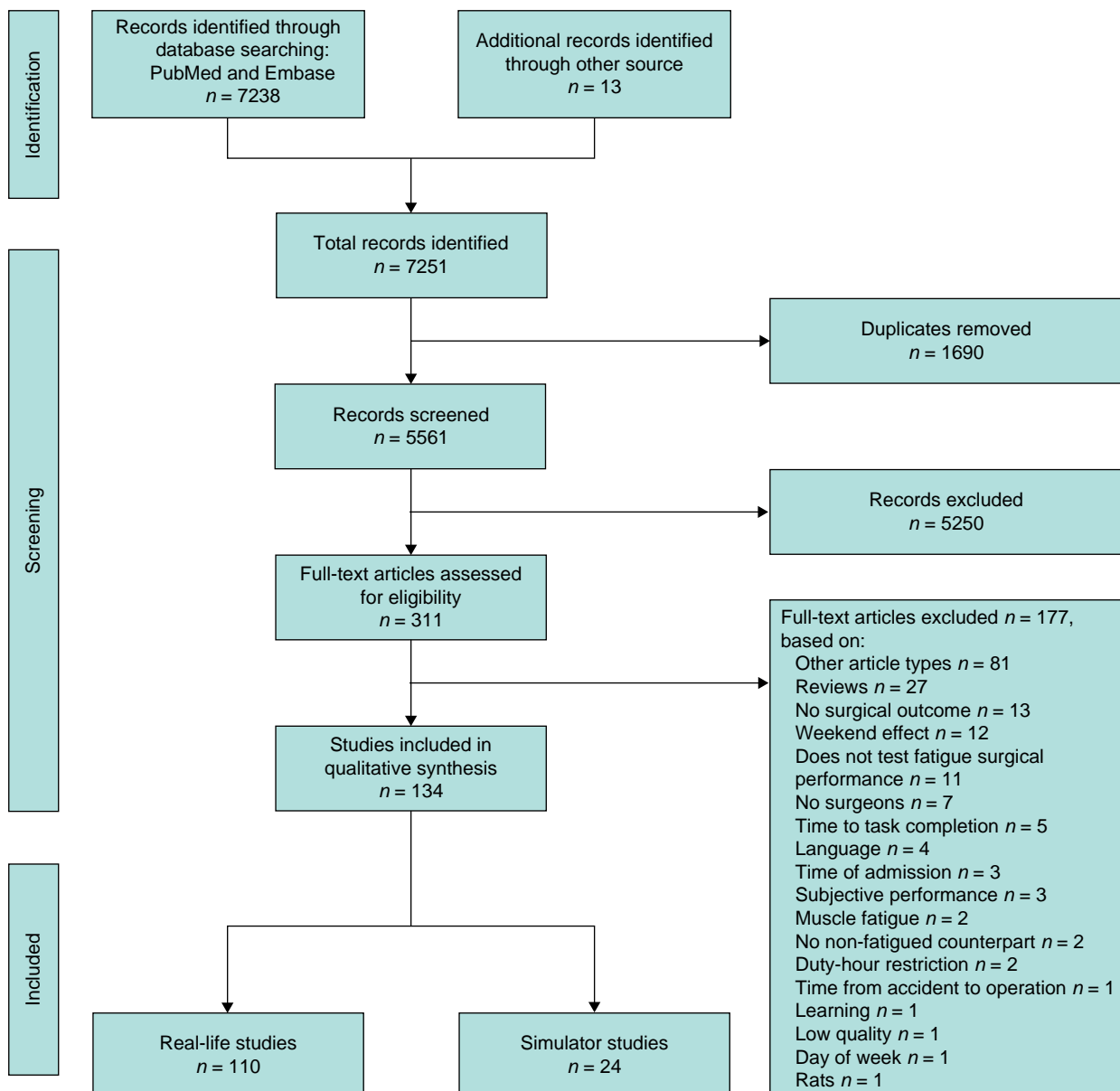


Fig. 1 PRISMA flow chart

quality of RCTs. The Newcastle–Ottawa scale was used to assess the quality of cohort studies¹³. ‘Poor-quality’ cohort studies or studies scoring at high risk of bias in any one of the six bias domains were excluded.

Data synthesis and analysis

Due to the large heterogeneity of outcome measures and study designs, a meta-analysis was not feasible. Instead, a narrative approach was used to summarize findings. Studies were categorized as either ‘real-life studies’ or ‘simulator studies’ to allow for comparison. See the [Supplementary Methods](#).

Results

Out of the 7251 studies identified, a total of 110 real-life studies and 24 simulator studies were included in the analysis (Fig. 1). Only one RCT was included in the analysis, scoring a low risk-of-bias for each item¹⁴.

Measures of fatigue

Studies used a large variety of measures to establish fatigue in surgeons. See [Table S1](#) and [Table S2](#). In the majority of real-life studies (86 studies), fatigue was assessed based on the time of day of the operation (for example daytime or night-time). Conversely, most simulator studies (17 studies) assessed fatigue in relation to the moment within the on-call shift that the measurement was taken. In contrast to any of the real-life studies, 13 simulator studies utilized validated questionnaires for a more direct measure of fatigue. Additionally, two simulator studies incorporated psychophysiological measurements, such as pupillography, to assess fatigue.

Surgical outcome in simulator studies

A total of 24 simulator studies, encompassing 460 simulation procedures, were included in the analysis. See [Table S3](#).

Simulation procedures consisted of either isolated laparoscopic proficiency tasks or simulated surgical operations. Among these studies, 18 studies solely assessed surgical outcome by psychomotor function, whereas 6 studies considered both psychomotor and cognitive functions.

A total of 11 simulator studies (46%) reported negative effects of fatigue on surgical performance. Of these, seven studies reported a decline in one or more psychomotor metrics when fatigue was present, whereas four studies reported a decline in both cognitive and psychomotor metrics, with three of them indicating a more pronounced effect on cognitive performance than on psychomotor performance, as evidenced by significant increases in cognitive errors and impaired performance in tasks requiring cognitive dominance.

Interestingly, one simulator study reported improved cognitive performance when fatigue was present¹⁵. According to Schlosser et al.¹⁵, this improvement may be rooted in coping strategies and the motivation levels of participants.

Overall, 50% of simulator studies (12 studies) reported no discernible effect on surgical outcome when fatigue was present.

Surgical outcome in real-life studies

A total of 110 real-life studies, encompassing a total of 1 683 613 surgical procedures, were included in the analysis. See [Table S4](#).

Overall, 35.4% of real-life studies (39 studies) reported a deterioration in surgical outcome when fatigue was present, 8.2% of real-life studies (9 studies) reported an improvement in surgical outcome when fatigue was present, 8.2% of real-life studies (9 studies) reported a mixed effect, that is both a deterioration and improvement, when fatigue was present, and 48.2% of real-life studies (53 studies) reported no discernible effect on surgical outcome when fatigue was present.

These studies employed a total of 1068 different patient outcome measures, which could be categorized into 10 distinct groups (Fig. 2). When translating the effects to the patient

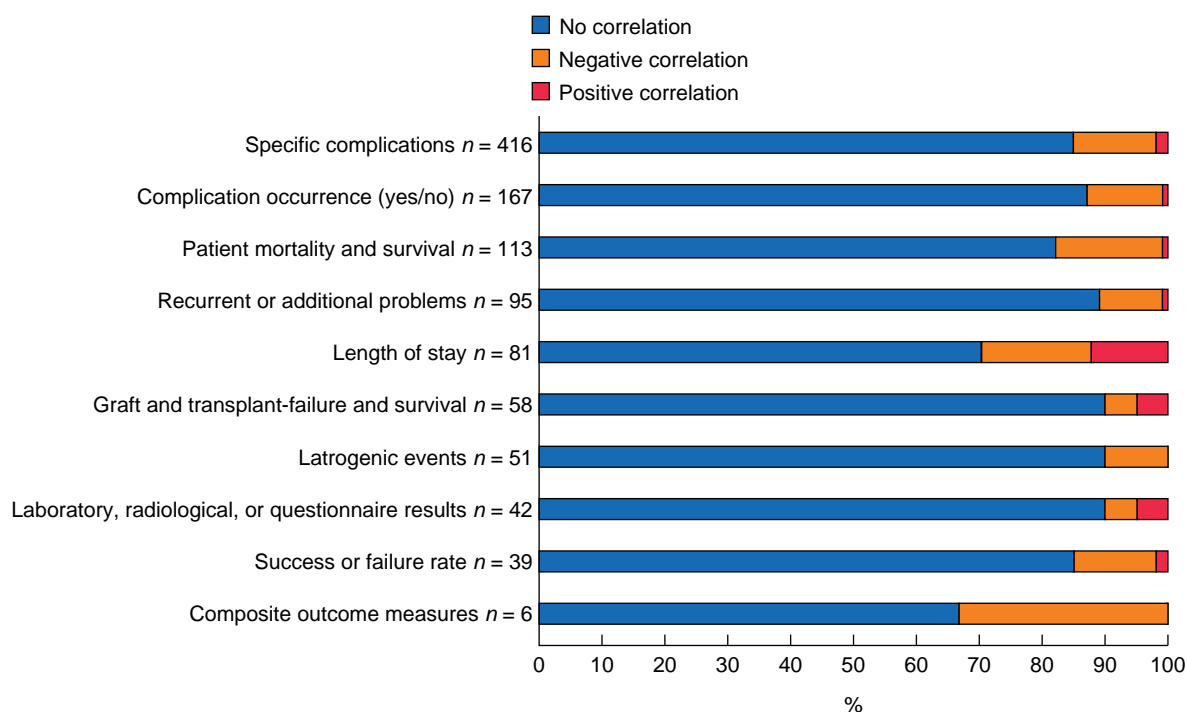


Fig. 2 Percentage of outcome measures negatively affected, positively affected, or unaffected by fatigue per category

outcome measures, the negative effect of fatigue was present in only 12.5% of all patient outcome measures (134 patient outcome measures), primarily in the category of 'specific complication mentioned' (55 patient outcome measures). See [Table S5](#).

Discussion

This systematic review assessed the impact of fatigue on surgical proficiency or outcomes, both in real-life and simulator studies. In simulator studies, fatigue negatively affected surgical performance in almost half of the studies. In contrast, among real-life studies, only one-third of studies showed a deterioration in patient outcomes when fatigue was present. This effect in real-life studies was limited to a subset of the wide-ranging spectrum of patient outcome measures evaluated. However, in the case of real-life studies, a wide range of findings was found and the evidence supporting these effects remains inconclusive.

Almost half of the simulator studies reported a negative effect of fatigue on surgical performance, impairing both cognitive function and psychomotor performance. The majority of identified simulator studies solely focused on psychomotor performance only, despite the known importance of cognitive function performance¹⁶. Studies reporting both skills showed a more profound decline in cognitive performance^{12,17–19}. Many surgical psychomotor skills are highly automatized and maintained during fatigue, for example a surgeon can adequately tie a knot even when fatigued, yet this is not the case for more cognitively demanding tasks^{20,21}. While routine operations may be competently completed when a surgeon is fatigued, decision-making in complex cases can never be automated and will always require effort and attention²². This aligns with numerous studies in other domains^{23,24}.

In real-life studies, a wide range of findings emerged. While some studies assert a negative effect of fatigue, the majority of studies do not report a significant difference in patient outcomes between fatigued and non-fatigued surgeons. Notably, the negative effects observed in the studies are often limited to a very small fraction of the large number and breadth of outcome measures employed.

The negative effects of fatigue, found in half of the simulator studies, seem to diminish in real-life studies, yet a lack of substantive knowledge on fatigue within study designs and flawed methodologies in the real-life studies limit the confidence that this is a true difference. For example, all of the real-life studies and half of the simulator studies lack direct measures of fatigue, such as validated fatigue questionnaires. In most of the real-life studies, the time of day of the operation is used as a proxy for fatigue, which assumes, without validation, that surgeons experience fatigue during night shifts²⁵. Furthermore, in real-life surgery, outcomes rely heavily on crude indicators of patient outcome, such as patient mortality or number of readmissions, which are impacted by multiple other factors and are typically analysed at an aggregated level rather than at the individual surgeon level. In contrast, simulator studies directly examine a surgeon's cognitive or psychomotor performance. By incorporating more specific behavioural outcome measures within otherwise controlled environments, simulator studies have the potential to detect more subtle differences between fatigued and non-fatigued surgeons. Finally, while some real-life studies report a significant association between surgeon fatigue and patient outcome, it is not possible to establish a direct causal relationship based solely on these findings given other potential variables impacting patient outcome, for example differences between and within surgical teams and in perioperative care. All such differences can influence outcomes,

leading to different results, even among patients with similar baseline conditions, regardless of a surgeon's state of fatigue. Consequently, while the aggregated measures used in these studies may provide a comprehensive overview of the collective surgical team's performance and its broader context, they fail to accurately represent an individual surgeon's contributions.

Fatigue emerges from prolonged engagement in a demanding task within a context of limited control. In high-stake situations (that is in real-life studies), individuals naturally strive to uphold performance at a satisfactory level by channeling increased effort²⁶. High intrinsic motivation can counteract the fatigue-induced impact on performance^{15,27–30}. However, while sustained high effort might temporarily mitigate the effects of fatigue, it can culminate in the experience of strain, a state associated with burnout³¹. Thus, surgical performance may be protected, but at a cost to the surgeon³².

Given the current methodological shortcomings, accurately assessing the true impact of fatigue on surgeon performance remains challenging. The findings of the present study support the need for future research to incorporate kinematic measures by means of instrument position trackers and accelerometers into real-life surgery^{33,34}. Additionally, integrating validated fatigue detection methods is imperative. Approaches, such as electrocardiography and eye metrics, which have already demonstrated effectiveness in detecting fatigue within the aviation industry, could prove valuable^{35,36}. By combining these kinematic and fatigue measurement approaches, more insight may be gained into the real effects of fatigue on surgical performance and also patient outcome. Furthermore, future studies should embrace a broader lens, encompassing team performance rather than focusing solely on the individual surgeon. Such an approach will enrich the understanding of the intricate interplay between various elements within the surgical environment.

This study has some inherent limitations. First, there was large heterogeneity in fatigue measures amongst the studies involved, as well as statistical analyses applied, complicating comparisons between studies. Furthermore, as almost all studies focusing on after-hours or sleep deprivation mentioned fatigue, it was assumed that the goal of these studies was to examine the effects of fatigue.

To conclude, almost half of the simulator studies and one-third of the real-life studies reported negative effects of fatigue, highlighting a significant patient safety issue. While in real-life studies, a difference in motivation might protect task performance, the current fatigue and outcome measures used, especially in real-life studies, are insufficient to accurately determine whether fatigue truly impacts surgical outcomes. Further research should include more direct measures of a surgeon's performance in real-life studies, measures of the effects on the whole surgical team, and a core clinical outcome set for surgical outcomes.

Funding

The authors have no funding to declare.

Acknowledgements

The authors would like to acknowledge Sjoukje van der Werf (information specialist at the Central Medical Library, University Medical Centre Groningen, Groningen, The Netherlands) for her support in the creation of the search strategy of this systematic review, as well as Neda Ansari (a medical student at the University of Groningen, Groningen, The Netherlands) for her contribution to this study.

Author contributions

Iris M. Reijmerink (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing—original draft, Writing—review & editing), Maarten J. van der Laan (Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Supervision, Writing—review & editing), J. K. Götz Wietasch (Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Supervision, Writing—review & editing), Lotty Hooft (Investigation, Supervision, Writing—review & editing), and Fokie Cnossen (Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Supervision, Writing—review & editing)

Disclosure

The authors declare no conflict of interest.

Supplementary material

Supplementary material is available at *BJS* online.

Data availability

The data that support the findings of this study are derived from public domain sources and are openly available within MEDLINE and Embase.

References

- Janhofer DE, Lakhiani C, Song DH. Addressing surgeon fatigue: current understanding and strategies for mitigation. *Plast Reconstr Surg* 2019;**144**:693e–699e
- Wetzel CM, Kneebone RL, Woloshynowych M, Nestel D, Moorthy K, Kidd J et al. The effects of stress on surgical performance. *Am J Surg* 2006;**191**:5–10
- Dawson D, Reid K. Fatigue, alcohol and performance impairment. *Nature* 1997;**388**:235
- Head J, Tenan MS, Tweedell AJ, LaFiandra ME, Morelli F, Wilson KM et al. Prior mental fatigue impairs marksmanship decision performance. *Front Physiol* 2017;**8**:680
- Bray RM, Camlin CS, Fairbank JA, Duntelman GH, Wheelless SC. The effects of stress on job functioning of military men and women. *Armed Forces Soc* 2001;**27**:397–417
- Armentrout JJ, Holland DA, O'Toole KJ, Ercoline WR. Fatigue and related human factors in the near crash of a large military aircraft. *Aviat Space Environ Med* 2006;**77**:963–970
- Gawande AA, Zinner MJ, Studdert DM, Brennan TA. Analysis of errors reported by surgeons at three teaching hospitals. *Surgery* 2003;**133**:614–621
- Ellman PI, Kron IL, Alvis JS, Tache-Leon C, Maxey TS, Reece TB et al. Acute sleep deprivation in the thoracic surgical resident does not affect operative outcomes. *Ann Thorac Surg* 2005;**80**:60–64; discussion 64–65
- McCormick F, Kadzielski J, Landrigan CP, Evans B, Herndon JH, Rubash HE. Surgeon fatigue: a prospective analysis of the incidence, risk, and intervals of predicted fatigue-related impairment in residents. *Arch Surg* 2012;**147**:430–435
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic Rev* 2021;**10**:89
- Sandblom G, Sevonius D, Staël von Holstein C. Impact of operative time and surgeon satisfaction on the long-term outcome of hernia repair. *Hernia* 2009;**13**:581–583
- Kahol K, Leyba MJ, Deka M, Deka V, Mayes S, Smith M et al. Effect of fatigue on psychomotor and cognitive skills. *Am J Surg* 2008;**195**:195–204
- Abou-Setta A, Beaupre L, Jones C, Rashed S, Hamm M, Sadowski C. Appendix 1, Newcastle-Ottawa scale assessment of cohort studies. In: *Pain Management Interventions for Hip Fracture*. Rockville, MD: Agency for Healthcare Research and Quality (US); 2011. <https://www.ncbi.nlm.nih.gov/books/NBK56664/>
- Uchal M, Tjugum J, Martinsen E, Qiu X, Bergamaschi R. The impact of sleep deprivation on product quality and procedure effectiveness in a laparoscopic physical simulator: a randomized controlled trial. *Am J Surg* 2005;**189**:753–757
- Schlosser K, Maschuw K, Kupietz E, Weyers P, Schneider R, Rothmund M et al. Call-associated acute fatigue in surgical residents—subjective perception or objective fact? A cross-sectional observational study to examine the influence of fatigue on surgical performance. *World J Surg* 2012;**36**:2276–2287
- Sahar Y, Wagner M, Barel A, Shoval S. Stress-adaptive training: an adaptive psychomotor training according to stress measured by grip force. *Sensors (Basel)* 2022;**22**:8368
- Brandenberger J, Kahol K, Feinstein AJ, Ashby A, Smith M, Ferrara JJ. Effects of duty hours and time of day on surgery resident proficiency. *Am J Surg* 2010;**200**:814–818; discussion 818–819
- Gerdes J, Kahol K, Smith M, Leyba MJ, Ferrara JJ. Jack Barney award: the effect of fatigue on cognitive and psychomotor skills of trauma residents and attending surgeons. *Am J Surg* 2008;**196**:813–819; discussion 819–820
- Kahol K, Smith M, Brandenberger J, Ashby A, Ferrara JJ. Impact of fatigue on neurophysiologic measures of surgical residents. *J Am Coll Surg* 2011;**213**:29–34; discussion 34–36
- Muffly TM, Espallat-Rijo LM, Edwards AM, Horton A. Operating room fatigue: is your twentieth surgical knot as strong as your first? *J Surg Educ* 2012;**69**:215–217
- Fox S, Kotelba A. Variational principle of least psychomotor action: modelling effects on action from disturbances in psychomotor work involving human, cyborg, and robot workers. *Entropy (Basel)* 2019;**21**:543
- Dwyer CP. An evaluative review of barriers to critical thinking in educational and real-world settings. *J Intell* 2023;**11**:105
- Hamann A, Carstengerdes N. Assessing the development of mental fatigue during simulated flights with concurrent EEG-fNIRS measurement. *Sci Rep* 2023;**13**:4738
- Lieberman HR, Bathalon GP, Falco CM, Kramer FM, Morgan CA, Niro P. Severe decrements in cognition function and mood induced by sleep loss, heat, dehydration, and undernutrition during simulated combat. *Biol Psychiatry* 2005;**57**:422–429
- van de Ven HA, Hulsege G, Zoomer T, de Korte EM, Burdorf A, Oude Hengel KM. The acute effects of working time patterns on fatigue and sleep quality using daily measurements of 6195 observations among 223 shift workers. *Scand J Work Environ Health* 2021;**47**:446–455
- Hockey G. A motivational control theory of cognitive fatigue. In: Ackerman PL (ed.), *Cognitive Fatigue: Multidisciplinary Perspectives on Current Research and Future Applications*. Washington, DC: American Psychological Association, 2011;167–187
- Herlambang MB, Taatgen NA, Cnossen F. The role of motivation as a factor in mental fatigue. *Hum Factors* 2019;**61**:1171–1185
- Boksem MAS, Meijman TF, Lorist MM. Mental fatigue, motivation and action monitoring. *Biol Psychol* 2006;**72**:123–132

29. Barte JCM, Nieuwenhuys A, Geurts SAE, Kompier MAJ. Motivation counteracts fatigue-induced performance decrements in soccer passing performance. *J Sports Sci* 2019;**37**:1189–1196
30. Herlambang MB, Cnossen F, Taatgen NA. The effects of intrinsic motivation on mental fatigue. *PLoS One* 2021;**16**:e0243754
31. Guan S, Xiaerfuding X, Ning L, Lian Y, Jiang Y, Liu J et al. Effect of job strain on job burnout, mental fatigue and chronic diseases among civil servants in the Xinjiang Uygur autonomous region of China. *Int J Environ Res Public Health* 2017;**14**:872
32. Demerouti E, Velduis W, Coombes C, Hunter R. Burnout among pilots: psychosocial factors related to happiness and performance at simulator training. *Ergonomics* 2019;**62**:233–245
33. Ghasemloonia A, Maddahi Y, Zareinia K, Lama S, Dort JC, Sutherland GR. Surgical skill assessment using motion quality and smoothness. *J Surg Educ* 2017;**74**:295–305
34. Wolf A, Ueda K. Contribution of eye-tracking to study cognitive impairments among clinical populations. *Front Psychol* 2021;**12**:590986
35. Wingelaar-Jagt YQ, Wingelaar TT, Riedel WJ, Ramaekers JG. Fatigue in aviation: safety risks, preventive strategies and pharmacological interventions. *Front Physiol* 2021;**12**:712628
36. Pan T, Wang H, Si H, Li Y, Shang L. Identification of pilots' fatigue status based on electrocardiogram signals. *Sensors (Basel)* 2021;**21**:3003