

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (http://bmjopen.bmj.com).

If you have any questions on BMJ Open's open peer review process please email <a href="mailto:info.bmjopen@bmj.com">info.bmjopen@bmj.com</a>

# **BMJ Open**

# The impact of fatigue and sleep deprivation on physician and patient outcomes: A systematic review

Journal:	BMJ Open
Manuscript ID	bmjopen-2018-021967
Article Type:	Research
Date Submitted by the Author:	26-Jan-2018
Complete List of Authors:	Gates, Michelle; University of Alberta, Pediatrics Wingert, Aireen; University of Alberta, Pediatrics Featherstone, Robin; University of Alberta, Pediatrics Samuels, Charles; Centre for Sleep and Human Performance Simon, Christopher; Canadian Medical Association Dyson, Michele; University of Alberta, Pediatrics
Keywords:	Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, SLEEP MEDICINE
	·

SCHOLARONE™ Manuscripts

The impact of fatigue and sleep deprivation on physician and patient outcomes: A systematic review

Michelle Gates, Aireen Wingert, Robin Featherstone, Charles Samuels, Christopher Simon, Michele P Dyson

**Michelle Gates**, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-486C, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Project coordinator

**Aireen Wingert**, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-490B, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Project coordinator

**Robin Featherstone**, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-486D, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Research librarian

**Charles Samuels**, Centre for Sleep and Human Performance, Suite 106, 51 Sunpark Drive SE, Calgary, Alberta, Canada, T2X 3V4
Physician and medical director

**Christopher Simon**, Canadian Medical Association, 1867 Alta Vista Drive, Ottawa, Ontario, Canada, K1G 5W8

Advisor, physician health

Michele P Dyson, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-474, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Assistant professor

Correspondence to: MP Dyson mdyson@ualberta.ca

Word count (main text): 4,853

**Objectives:** For physicians in independent practice, we aimed to synthesize evidence relevant to the Canadian context on two questions: (1) what are the impacts of sleep loss and fatigue on physician health and performance, and patient safety; (2) what is the effectiveness of interventions that target sleep loss and fatigue, in terms of physician and patient outcomes?

**Design:** We conducted a systematic review of online literature. Following a pilot phase, one reviewer independently selected studies by title and abstract; full texts were then reviewed in duplicate. One reviewer extracted data; another independently assessed a random 10% sample. Two reviewers assessed risk of bias. We synthesized results narratively.

**Data sources:** We searched Medline, Embase, PsycINFO, CINAHL and PubMed for published studies in April 2016; we updated the Medline search in November 2017. To locate unpublished studies, we searched Embase for conference proceedings since 2000, and hand-searched relevant meeting abstracts and association and foundation websites.

**Eligibility criteria for selecting studies:** English or French language primary research studies of any design published from 2000 to 2017 that examined the effect of fatigue or sleep-related exposures or interventions on any outcome among physicians in independent practice and their patients were eligible.

**Results:** We included 47 quantitative studies of variable quality. 28 studies showed associations between fatigue or sleep deprivation and physician health and well-being, specifically burnout, stress, adverse mental health outcomes, and reduced life satisfaction. 21 studies showed no impact on surgical performance, and mixed findings for psychomotor performance, work performance, and medical errors. Six cohort studies showed little to no effect on patient outcomes related to surgical or obstetric procedures.

**Conclusions:** Fatigue and sleep deprivation have detrimental effects on the health of physicians in independent practice. Due to numerous methodological shortfalls, the current body of evidence is inadequate to inform strong practice recommendations.

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

#### STRENGTHS AND LIMITATIONS OF THIS STUDY

- This rigorously conducted and transparently reported systematic review is the first to synthesize evidence on the effects of fatigue and sleep loss on physicians in independent practice.
- The review is timely, given recent calls for research into individual and organisational solutions for burnout, and an increased focus on physician health.
- While we have identified a diverse body of evidence, we could not draw definitive conclusions due to methodological weaknesses and heterogeneous outcome measures in the included studies.
- We have focused on evidence from high income countries; our findings may not be generalizable to other settings.

The working hours of physicians, which have been historically long and unpredictable, have been a topic of debate for many years.[1] Beginning in the late 1980s, evidence indicating that medical resident fatigue could negatively impact their cognitive functioning and performance, resulting in an increased risk of medical error, began to accumulate.[2] In response, by the early 2000s physicians' regulatory bodies worldwide began to take action toward restricting the work hours of medical residents and ensuring adequate time for recovery between shifts.[3-5] Since their implementation in the United States by the Accreditation Council for Graduate Medical Education (ACGME), the impact of work hour regulations has been widely researched. Still, evidence for impacts on patient care, resident training and wellbeing remains equivocal.[6-9] This is likely because work hours are only one of many contributors to fatigue and physician wellbeing. In fact, the ACGME has recently reversed the 2011 changes that limited resident work hours to 16 hours per shift and the requirement for 8 hours of time off between shifts. This decision was made in favour of promoting "flexibility" for residency training program work hours and scheduling.

The focus on medical trainees has left physicians in independent practice as a relatively neglected group in research and policy. In Canada, there is no concrete regulation on the hours or patterns in which physicians choose to work.[10] In the absence of clear policies, physicians trained under traditional systems (i.e., prior to resident work hour regulations) may find it difficult to work shorter hours or take more frequent breaks.[1] Indeed, more than 40% of practicing physicians in the United States work in excess of 80 hours per week.[11] While long work hours remain a cultural norm in medicine, in comparable high-risk industries (e.g., aviation), work patterns and work hours are tightly regulated.[12] The need for similar evidence-based policies in medicine has become a topic of increased interest. Exemplar of this, an evidence-based guideline for fatigue risk management in emergency medical services,[13] informed by a comprehensive set of systematic reviews, has recently been published. For physicians, it has been argued that there is a need to adapt healthcare systems and provide support in identifying the signs of fatigue and mitigating its risks.[1]

Besides potentially affecting patient outcomes, fatigue can impact the health and wellbeing of physicians themselves. Burnout, just one outcome related to fatigue, has been described as epidemic among physicians[14-16] and ultimately affects recruitment and retention of physicians both in community and acute care settings. While the effect of physician wellbeing on the sustainability of

healthcare systems has recently received increased attention,[17] evidence-based solutions to burnout remain relatively elusive.[18] What is clear, is that comprehensive organisational-level efforts are necessary to fully address the issue.[19] Research addressing the factors that influence burnout and overall physician wellness is needed to inform system- and individual-level strategies.[20, 21] To date, evidence of the effects of fatigue and the role of chronic sleep restriction on physicians in independent practice has not been synthesized, making it unclear what gaps in knowledge remain unaddressed.

Given this void, we undertook a systematic review of primary research relevant to the Canadian context, to examine the effects of fatigue and chronic sleep restriction on physicians in independent practice, and on interventions to combat these effects. Our review was guided by the following research questions: Among physicians in independent practice, (1) what are the impacts of fatigue and chronic sleep restriction on physician health, physician performance, and patient safety; and (2) what is the effectiveness of interventions that target fatigue and chronic sleep restriction loss, in terms of improving physician and patient outcomes?

## **METHODS**

# **Review conduct**

The conduct of this systematic review was guided by Cochrane standards.[22] The research team convened to plan the key research questions and methodology but did not register a formal protocol. The findings are reported in adherence with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) Statement.[23] Ethical approval was not required for this study.

# **Patient involvement**

Patients were not involved.

#### Literature search

An information specialist developed a search strategy that included concepts related to physicians, fatigue and sleep. On 13 April 2016 we searched the following online databases with coverage in the biomedical sciences and psychology: Medline, Embase, PsycINFO, CINAHL and PubMed, limited to English and French language articles published from 2000 to 2016 (Medline search updated in November 2017). Though fatigue among physicians is not a new phenomenon,[2] we limited our search to articles published post-2000 to include studies relevant to current physician practice. Work hour

limitations have existed in European countries since 1993, but implementation in the United States (2003)[5] and Canada (2013) for residents is more recent.[24] We aimed to include studies published in this era of increased awareness about the potential impacts of long work hours. To locate unpublished studies, we searched Embase for conference proceedings since 2000 and hand-searched meeting abstracts of the Canadian Conference on Physician Health and the International Conference on Physician Health (2012 to 2016). We also searched the following association and foundation websites: American Medical Association, Australian Medical Association, British Medical Association, Canadian Medical Association, European Medical Association, National Sleep Foundation, Ontario Medical Association and the World Medical Association. The complete search strategy undertaken is reported in Supplementary file 1.

#### **Inclusion criteria**

Primary studies (quantitative or qualitative) of fatigue- or sleep-related exposures or interventions among physicians in independent practice were eligible for inclusion. We included physicians practicing in any medical specialty and in any healthcare setting within a high income country, [25] to identify practices comparable to the Canadian setting. Studies including physicians-in-training were included only if data for physicians in independent practice could be isolated. Exposures of interest included fatigue, sleep restriction, or sleepiness. We also included studies of any intervention that aimed to reduce fatigue or sleep restriction with any comparator (or no comparator). All reported outcomes, measured at any time, were eligible for inclusion.

We excluded commentaries, letters, editorials and dissertations. Systematic reviews, health technology assessments, economic evaluations and practice guidelines were excluded, although the reference lists were scanned for potential primary studies for inclusion. Studies that focused solely on physicians-intraining (e.g., trainees, residents, fellows, interns, medical students, junior doctors, registrars) were ineligible. To maintain the focused scope of the review, we excluded work hours, work load, and any other exposure or intervention that was indirectly related to fatigue or sleep restriction.

## Study selection

Two reviewers piloted the selection criteria for title and abstract screening in duplicate on 300 records. Following the pilot phase, the reviewers applied the criteria independently to the remaining records. Then, we retrieved all records classified as "include" or "unsure" and the two reviewers assessed their

full text for eligibility, in duplicate. Disagreements during the full-text screening phase were resolved by discussion or the involvement of a third reviewer, when needed.

#### **Data extraction**

Reviewers used a standardized form to extract data in Microsoft Office Excel (v. 2016, Microsoft Corporation, Redmond, WA). One reviewer independently extracted data from each included study and a second reviewer verified a random 10% sample. Since no major errors or omissions were noted, we did not undertake further verification.

We extracted the following data: country of publication; funding source; study design; inclusion and exclusion criteria; population characteristics (i.e., sample size, age and gender distribution, physician specialty); setting (i.e., physician workplace, urban or rural); exposure or intervention; definition of fatigue or sleep loss; sleep and fatigue scales used and timing of measurement; comparators (if applicable); and outcomes.

## Risk of bias appraisal

Two reviewers independently assessed the risk of bias in each included study using standard tools. Disagreements were resolved via discussion or by consulting a third reviewer. We used the Cochrane Risk of Bias tool[22] to assess randomised controlled trials across seven domains: sequence generation; allocation concealment; blinding of participants and personnel; blinding of outcome assessors; incomplete outcome reporting; selective outcome reporting; and other sources of bias. Adapted versions of the tool developed by the Effective Practice and Organization of Care group[26] were used to assess before-after and time series studies. We used the Newcastle-Ottawa Quality Assessment Scale[27] to appraise cohort studies across three domains: sample selection; comparability; and outcome assessment. We adapted the scale to assess cross-sectional studies and the one non-comparative study.

#### **Evidence synthesis**

We found insufficient homogeneity in populations, exposures or interventions, and outcomes to pool the data via meta-analysis. We have presented the findings narratively and in summary tables.

#### **RESULTS**

We identified 16,083 unique records via the database searches, 56 grey literature sources, and 15 additional records in reference lists of systematic reviews. We excluded 15,016 citations by title and abstract, and another 1,091 by full text. Forty-seven studies[28-74] were eligible for inclusion. Figure 1 shows the flow of studies through the selection process.

# **Included study characteristics**

A summary of the study characteristics is provided in Table 1. Supplementary file 2 presents descriptive information for each included study. There were 45 observational studies [28-34, 36-69, 71-74] and two intervention studies. [35, 70] All studies were quantitative. Nearly half (n = 20/47, 43%) of the studies took place in North America, [29, 31-33, 36, 40, 43, 52-55, 57, 58, 60-63, 67, 72, 74] and slightly more than one-third (n = 16/47, 34%) in Europe. [28, 30, 34, 35, 37, 41, 42, 45-48, 56, 59, 68, 70, 71]

**Table 1.** Summary characteristics of the included studies

Study characteristics	n	%	Physician characteristics	n	%	Exposures, interventions and	n	%
						outcomes		
Study design			Gender			Exposures (observational) <sup>a</sup>	45	96
Cross-sectional	34	72	Reported <sup>b</sup>	38	81	Fatigue	15	32
Cohort	6	13	>50% male	30	79	Sleep deprivation	37	79
Before-after	3	6	Age			Overnight or extended shifts	18	38
RCT	2	4	Reported <sup>b</sup>	38	81	Interventions (experimental)	2	4
Time series	1	2	Range (years)	20 to	>70	Outcomes		
Non-comparative	1	2	Specialty area <sup>c</sup>			Physician health and wellbeing	28	60
Region and country			Surgeons	13	28	Work and life satisfaction	9	19
North America	20	43	Anesthesiologists	10	21	Burnout	7	15
US	15	32	Generalists	7	15	Stress	8	17
Canada	4	9	ED or ICU physicians	3	6	Mental health and wellbeing	7	15
Canada, US & Mexico	1	2	Oncologists	2	4	Other health-related outcomes	5	11
Europe	16	34	Obstetrician-gynecologists	1	2	Physician performance, risk of error	21	45
France	4	9	Mixed groups	14	30	Psychomotor performance	7	15
Finland	3	6	Work setting <sup>d</sup>			Work ability and quality of care	5	11
Spain	2	4	Hospitals	37	78	Incidence of medical errors	5	11
Austria	2	4	Private practice	13	28	Surgical efficiency, effectiveness	5	11
Norway	2	4	Primary care centres, outpatient clinics	7	15	Patient outcomes	6	13
Denmark	1	2	Academic practice, training programs	5	11			
Germany	1	2	Other (e.g., industry, military)	11	23			
Malta	1	2	Not reported	3	6			
Japan	4	9	Urban or rural			7/		
Australia	2	4	Reported <sup>b</sup>	16	34			
Israel	2	4	Urban	12	75			
New Zealand	2	4	Rural	2	13			
United Kingdom	1	2	Mixed	2	13			

ED: emergency department; ICU: intensive care unit; RCT: randomised controlled trial; US: United States of America

<sup>&</sup>lt;sup>a</sup>Exposures that have been directly related to an outcome. Some studies included multiple exposures.

<sup>&</sup>lt;sup>b</sup>Percentages presented using the total number of studies where the outcome was reported as the denominator.

<sup>&</sup>lt;sup>c</sup>Anesthesiologists include physician anesthetists; generalists include primary care physicians, internists, and general practitioners; mixed groups refers to studies including more than one physician group or specialty (usually large-scale surveys). In some studies, multiple distinct groups were represented.

<sup>&</sup>lt;sup>d</sup>As defined by the authors. Values for the settings will exceed 100% because studies may occur in more than one setting.

Fifteen (32%) studies reported on fatigue exposure, [29, 35, 40, 43, 52, 58-66, 68] while others (n = 37/47, 79%) reported on sleep deprivation or reduced sleep quality. [28, 30-42, 44-51, 53-57, 59, 62, 66, 67, 69-74] A few (n = 5/47, 11%) reported on both. [35, 40, 59, 62, 66] In some cases (n = 18/47, 38%), fatigue or sleep loss were related to overnight work or long on-call shifts. [28, 31, 33, 35, 36, 38, 40, 41, 45, 46, 48, 53-55, 58, 70, 72, 74] Measured outcomes varied widely and were ultimately organised into physician physical and mental health, physician performance and risk of error, and patient outcomes.

# Risk of bias appraisal

The overall quality of the body of research was poor; 72% (n = 34/47) of studies were rated at unclear or high risk of bias. Of the two randomised controlled trials, one was rated as unclear overall risk of bias[70] and one as high risk.[35] All cohort studies were at low risk of bias (mean score: 8.4/9, range: 8-9).[33, 36, 53, 58, 72, 74] All of the before-after studies were rated as high risk of bias.[28, 40, 45] The single time series study was assessed at high risk of bias.[46] The cross-sectional studies varied in performance (mean score: 3.0/5, range: 1-4); only one-third (n = 12/34, 35%) were at low risk of bias.[34, 37, 39, 42, 47, 54, 55, 64, 65, 67, 69, 71] The one non-comparative study was at unclear risk of bias.[38] Detailed assessments for each study are shown in Supplementary file 3.

# Physician health and wellbeing outcomes

Twenty-eight studies reported on physician health and wellbeing-related outcomes, [29, 30, 32, 34, 35, 37, 41-43, 45-52, 55, 57, 59, 62, 63, 65-67, 69, 71, 73] including burnout (n = 7), stress (n = 8), mental health and wellbeing (n = 7), life and job satisfaction (n = 9) and other markers of health (n = 5) (Supplementary file 4).

Seven cross-sectional studies demonstrated links between sleep deprivation and burnout among surgeons, [49, 57] anesthesiologists, [34] generalists [71] and other mixed groups. [65, 67, 69] Two studies reported on surgeons; the larger (n = 2,564, low risk of bias) study of neurosurgeons showed increased odds of burnout with sleep deprivation (OR 0.84, 95% CI 0.75-0.94, P = 0.002). [49] Among anesthesiologists one study (n = 565, low risk of bias) indicated that burnout was more prevalent among the sleep-deprived (47.6% vs. 16.3%, P < 0.001). [34] In one small (n = 11) study of generalists, those with burnout had poorer Pittsburgh Sleep Quality Index scores (7.24 $\pm$ 4.17 vs. 2.72 $\pm$ 2.22, P < 0.001). [71] In the two larger studies of mixed physician groups (low risk of bias), burnout was more prevalent among those who were sleep deprived (39.6% vs. 26.4%, P < 0.05), [67] and physical fatigue was correlated with burnout (r = 0.88, P < 0.05). [65]

Seven observational studies of varying methodological quality[29, 41, 42, 45, 47, 57, 59] and one intervention study at high risk of bias[35] reported on stress outcomes among surgeons,[57] anesthesiologists,[45, 47] emergency physicians,[35, 59] internal medicine physicians,[41] and mixed groups.[29, 42] In a small sample (n=20) of internal medicine physicians, a 24-hour call shift had no effect on biochemical or physiological stress parameters, except levels of thyroid stimulating hormone, which was higher post-shift (P = 0.049, data not reported).[41] The remaining observational studies suggested that there was a link between sleep deprivation or fatigue and stress. The one study of orthopedic surgeons (n = 264, high risk of bias) showed that sleep deprivation and psychological distress were correlated (data not reported, P < 0.001).[57] The two reports on anesthesiologists were of varied quality; the larger (n = 328) study that was at low risk of bias showed that stress symptoms were predicted by sleep deprivation ( $\beta = -0.269$ , P < 0.001).[47] Among the two studies reporting on mixed groups of physicians, the larger (n = 1,541, low risk of bias) study showed an association between sleep problems and psychological distress ( $\beta = 0.18$ , P < 0.001).[42] One RCT assessed the impact of sleep deprivation from shift work, showing that stress among emergency physicians (n = 17) was higher following the shift as compared to a control day (data not reported, P < 0.05).[35]

Seven cross-sectional studies of varying methodological quality reported on aspects of mental health including addiction or substance misuse, [30, 48, 66] depression, [73] thoughts of suicide, [47] mood disturbance [55, 66] and overall wellbeing. [62] One study, [48] which was at high risk of bias, showed no association between hours of sleep when on call and hazardous drinking behaviours. Meanwhile, the six other studies all showed deleterious effects of sleep deprivation and fatigue on mental health. Three

studies reported on anesthetists, [30, 47, 55] with two large surveys showing increased odds of tobacco (OR 1.42, 95% CI 1.04-1.94) and tranquilizer/hypnotics (OR 3.26, 95% CI 2.12-5.02) dependency being predicted by sleep deprivation, [30] and sleep disturbance being associated with thoughts of suicide (P = 0.009). [47] A small study (n = 21) showed greater mood disturbance following a 17-hour night shift than a usual day (Profile of Mood States score 42.57±15.26 vs. 70.90±6.91, P < 0.001). [55] Among oncologists (n = 241), overall wellbeing was predicted by lower levels of fatigue after controlling for personal and professional characteristics (P = 0.002). [62] A large (P = 0.002). [63] In open-ended questions, senior physicians in one study (unclear risk of bias) attributed the development of mental illness to tiredness and stress at work. [66]

Nine cross-sectional studies of varying methodological quality reported on outcomes related to job satisfaction, [37, 42, 43, 50, 67, 69] life satisfaction [32, 57, 67] or work-life balance. [63, 67] All but two [37, 67] of these studies showed that sleep deprivation and fatigue were associated with reductions in satisfaction. The six studies that investigated job satisfaction were all at low risk of bias and generally included mixed groups of physicians [42, 67, 69]; one study reported on general practitioners, [50] and another on surgeons. [43] Three studies showed that reductions in sleep duration and/or quality [42, 43, 69] were associated with reduced job satisfaction. Meanwhile one showed no association between sleep deprivation and career satisfaction, [67] and another showed no relationship between earlier sleep disturbance and later job demands or job control. [37] A single study (n = 92) reporting on rural general practitioners indicated that frequent sleep disturbance predicted the intention to retire early (OR 2.91, 95% CI 1.11-7.6, P < 0.05). [50]

The three studies that reported on life satisfaction were of variable quality, but all demonstrated links between sleep deprivation or fatigue and reductions in life satisfaction.[32, 57, 67] Of two studies among mixed physician groups,[32, 67] the one larger (n = 840) study showed that sleep deprivation (less than 7 hours per day) was a predictor of reduced life satisfaction (OR 0.44, 95% CI 0.29-0.67, P < 0.05).[67] One study at high risk of bias reported on orthopedic surgeons (n = 264), showing that sleep deprivation was correlated with lower marital satisfaction (data not reported, P < 0.001).[57] Two large studies at low or unclear risk of bias reported on work-life balance.[63, 67] Among oncologists (n = 1,117), reduced satisfaction with work-life balance was predicted by high levels of fatigue, even when

controlling for personal and work-related factors and burnout (OR 0.489, 95% CI 0.337-0.710, P < 0.001).[63] Among a mixed group of physicians (n = 840, low risk of bias), sleep deprivation predicted a reduced perception of having balanced personal and professional commitments (OR 0.46, 95% CI 0.31-0.71,  $P \le 0.05$ ).[67]

Five cross sectional studies at high or unclear risk of bias[32, 51, 52, 66] and one time series study at high risk of bias[46] reported on other health-related outcomes. Among a mixed group of physicians (n = 180), one study at high risk of bias showed that Epworth Sleepiness Scale scores were higher among physicians who worried about having a car accident while driving home (7.0 vs. 5.4, P < 0.001).[32] Among generalists (n = 578), almost 1 in 10 (8.7%) admitted to falling asleep while driving due to fatigue.[52] Also among generalists (n = 92), those with frequent work-related sleeping problems were at increased odds of sickness presenteeism (OR 2.92, 95% CI 1.19-7.16, P = 0.02).[51] The one time series study concluded that a single 24-h shift did not cause major chronodisruption among anesthetists (n = 10).[46] Meanwhile, open-ended comments from a large sample (n = 3,550) of senior physicians suggests that they attributed the development of physical health problems to a lifestyle of sleep deprivation, poor eating habits and lack of exercise imposed by their jobs.[66]

## Physician performance and risk of errors

Twenty-one studies reported on physician performance and safety-related outcomes, [28, 31-33, 36, 38-42, 44, 45, 54, 56, 60, 61, 64, 66, 68, 70, 74] including surgical efficiency and effectiveness (n = 5), psychomotor performance (n = 7), work ability and quality of care (n = 5) and incidence of medical errors (n = 5) (Supplementary file 5).

Three cohort studies at low risk of bias,[33, 36, 74] one before-after study at high risk of bias[28] and one randomized controlled trial at high risk of bias[70] examined the effects of sleep deprivation from overnight work or extended shifts, during surgeries[33, 36, 74] or laparoscopic simulations.[28, 70] The cohort studies, which reported on 49,776 surgical procedures, found no adverse effects on any measure of surgical efficiency or effectiveness.[33, 36, 74] The small (n = 29) before-after study showed no impact of sleep deprivation from shift-work nor of sleep hours on performance on a laparoscopic simulation.[28] One small (n = 64) intervention study compared a 24-hour shift to a usual work day, also finding no detriment to performance on a laparoscopic simulation despite diminished sleep hours while working on-call.[70]

Two before-after studies at high risk of bias[40, 45] and five cross-sectional studies of variable methodological quality[31, 38, 41, 54, 56] reported on psychomotor performance outcomes among surgeons, [40] anesthesiologists, [31, 38, 45, 54] emergency physicians, [56] and internal medicine physicians.[41] Four studies[38, 40, 45, 54] showed an overall reduction in psychomotor performance in the fatigued state while the others had mixed results.[31, 56] Among a small group of surgeons (n = 9), performance on a virtual ring transfer task deteriorated after an on-call shift (data not reported, P < 0.05).[40] The four studies among anesthetists reported mixed findings. One small (n = 11) before-after study showed longer reaction times (690.8±73.4 vs. 746.5±113.7 milliseconds) and reduced concentration ability (26.4±23.5 vs. 56.3±23.0 on a 100-point scale, P = 0.007) following a 24-hour shift with sleep deprivation[45]; Two others found that sleep loss was associated with slower reaction times.[38, 54] Conversely, a small study (n = 11) found no effect of overnight shiftwork with sleep deprivation on any measure of psychomotor performance except Hopkin's Verbal Learning Test (t-score of  $48.6\pm7.6$  vs.  $41.5\pm9.9$ , P = 0.04).[31] Among emergency physicians (n = 18), one study (unclear risk of bias) those who were sleep deprived had a reduced performance on most but not all psychomotor tests, [56] while among internal medicine physicians (n = 20, low risk of bias), neurocognitive parameters did not seem to worsen post-call.[41]

Five cross-sectional studies of variable methodological quality reported on associations between sleep deprivation or fatigue and work ability or perceived performance, all among mixed groups of physicians.[32, 42, 60, 64, 66] The two large studies at low risk of bias showed that sleep problems and fatigue had a negative impact on physicians' work. [42, 64] Among 1,541 physicians in Finland, sleeping problems were inversely associated with scores on the Work Ability Index ( $\beta = -0.29$ , P < 0.001),[42] while a study of 890 physicians from Israel demonstrated that perceived quality of care was predicted by fatigue even after controlling for components of burnout ( $\beta = 0.17$ , P < 0.05).[64] Similarly, in one study, comments from senior physicians suggested that continual tiredness and exhaustion negatively affected their competence. [66] The two studies [32, 60] that were at high risk of bias had conflicting findings.

Five cross-sectional studies of variable methodological quality reported on associations between sleep deprivation, fatigue and self-reported medical errors among surgeons, [61] anesthesiologists [39] and mixed groups of physicians.[32, 44, 68] Two studies showed that sleep disturbance was associated with an increased risk of errors, [39, 44] while the findings of the other studies were mixed. [32, 61, 68] A large

(n = 7,905) study at unclear risk of bias showed that only 6.9% of surgeons reported fatigue as the most important contributor to medical errors. [61] Among anesthesiologists, a smaller study (n = 183) at low risk of bias showed that the risk of fatigue-related errors increased with more nights of work-related sleep disturbance (RR 1.25, 95% CI 1.06-1.49). [39] Two of the studies reporting on mixed groups of physicians had conflicting results, [32, 44] while another reported that physicians' opinions on the association between fatigue and prescribing errors differed by work setting. [68] One-third (34%) of community-based, 96% of hospital-based, and 8% of office-based physicians believed that there was a high or very high association between fatigue and prescribing errors (P < 0.05). [68]

## **Patient Outcomes**

Six large (n = 270 to 38,978) cohort studies at low risk of bias reported on patient outcomes, all related to surgical[33, 36, 53, 58, 72, 74] or obstetric[53] procedures (Supplementary file 6). In these studies, sleep deprivation or fatigue were typically defined as overnight work prior to a daytime procedure[36, 53, 58, 72, 74]; though two studies measured sleep hours[33] or 'sleep opportunity'.[53] Overall there appeared to be little[33] to no[36, 58, 72, 74] effect of sleep deprivation from overnight work on adverse patient outcomes such as operative complications, length of stay, and mortality. One study showed that nighttime work prior to a daytime procedure did not affect complication rates, but that shorter sleep opportunity increased the odds of operative (OR 2.70, 95% CI 1.13-6.48, P = 0.03) but not obstetric complications.[53]

# **DISCUSSION**

Fatigue and chronic sleep restriction are two potential drivers of reduced physician wellbeing[17, 19] that have thus far been understudied in physicians in independent practice. Burnout is becoming increasingly prevalent among physicians,[14-16] and recent research indicates that comprehensive individual- and system-level strategies are needed to address the problem.[6-9, 19, 21] We have systematically reviewed evidence from a heterogeneous array of available studies reporting on diverse outcomes related to physicians in independent practice and their patients. The included studies were often at high or unclear risk of bias, included small samples of physicians, and inconsistently measured and reported exposures and outcomes. As a result of these methodological shortcomings, the currently available evidence is inadequate to inform practice or policy recommendations.

In light of high rates of burnout, the ongoing dialogue about the need for a cultural shift in the practice of medicine[79, 80] is now more important than ever. Recognition of the potential effects of physician fatigue on patients, physicians, and healthcare systems as a whole must be emphasized at a systemic level, encouraging a shift in which the risks are viewed as unacceptable.[1, 20, 76] Likewise, although research to date has focused largely on individual-level approaches to address burnout, it is now clear that placing the burden of a system-level problem solely on the individual is unlikely to bring about significant and lasting change.[81] Recent research has highlighted physician burnout as a system-driven issue that will require corresponding national-scale multicomponent solutions.[1, 19, 77, 78] As such, in the past several years both the American and Canadian Medical Associations have developed policies and programs that address physician health.[77, 82] The Canadian Medical Association's new policy on physician health calls on broad stakeholder groups (e.g., policymakers, regional health authorities, governments) to take shared responsibility for the health of physicians and to make meaningful and concerted efforts towards promoting a healthy and sustainable workforce.[77]

Our systematic review indicates that the current evidence base is inadequate to inform decision-making. Correspondingly, a 2016 research summit on physician wellness and burnout outlined the need for timely, relevant and methodologically robust research to inform practice and policy.[21] We identified only two intervention studies, which supports the assertion that novel interventions with realistic budgets and timelines at both individual and organisation levels need to be tested.[21] The vast array of tools used by current studies to measure sleep, fatigue and various outcomes impedes evidence

synthesis. It will be important to make use of exiting validated measures[83-85] consistently in future research. Identifying outcomes of importance to physicians and their patients should be prioritized, such that these may be collected within intervention studies. Reporting these consistently will allow for effective synthesis of findings and reduce research waste.[86] Integrated knowledge translation strategies involving multiple stakeholder groups (e.g., physicians, patients, medical schools, physicians' associations and governing bodies, policymakers) may help to ensure that the research is relevant and facilitates decision-making.[87]

## **Strengths and Limitations**

Our systematic review is the first to synthesize evidence on the effects of fatigue and sleep loss on physicians in independent practice. The review is timely, given recent calls for research into individual and organisational solutions for burnout, [20, 21] and an increased focus on physician health. [76, 77] While we have identified a diverse body of evidence, we could not draw definitive conclusions due to methodological weaknesses and heterogeneous outcome measures in the included studies. The findings may have been influenced by publication bias, and may not be generalized to all settings, given our restriction to high income countries. Rigorously conducted and reported studies will be required to determine with confidence the potential impacts of fatigue on physicians and their patients, and to inform reasonable and sustainable solutions to the problem.

## **CONCLUSION**

The evidence synthesized in this review suggests some detrimental impacts of fatigue and sleep deprivation on physician health and wellbeing, and mixed evidence for potential impacts on performance and safety outcomes. The evidence overall did not indicate any impact on patient outcomes. Our overall confidence in the findings is low, owing to a body of research that is hindered by methodological weaknesses, including small sample sizes and inconsistent measurement of fatigue exposure and outcomes. Further methodologically robust research that includes consistent outcomes that are of interest to physicians and their patients is needed to inform strong practice recommendations and policy decisions.

#### **ACKNOWLEDGMENTS**

We would like to thank Dr. Meghan Sebastianski and Andrea Milne for assistance with screening, and Dr. Allison Gates for screening, data verification, and quality assessment of the French studies.

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide license to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material whereever it may be located; and, vi) license any third party to do any or all of the above.

## **COMPETING INTERESTS**

All authors have completed the ICMJE uniform disclosure form at <a href="www.icmje.org/coi\_disclosure.pdf">www.icmje.org/coi\_disclosure.pdf</a> and declare: Dr. Christopher Simon is employed by the Canadian Medical Association, who provided financial support for the research; there are no other relationships or activities that could appear to have influenced the submitted work.

# **CONTRIBUTOR STATEMENT**

All authors contributed to the conception and design of the project. MG and AW contributed to the acquisition, analysis and interpretation of the data, and drafted the manuscript. RF contributed to acquisition of data. CSa, CSi and MPD contributed to interpretation of data and revised the manuscript for important intellectual content. All authors approved the final version of the manuscript as submitted.

#### **FUNDING STATEMENT**

This research was supported by the Canadian Medical Association. The funder had no role in the study design; collection, analysis and interpretation of data; the writing of the report; and the decision to submit the article for publication.

# **ROLE OF FUNDERS**

Dr. Christopher Simon is employed by the Canadian Medical Association. The remaining authors are independent from the funders. The funders had no role in the study design; in the collection, analysis,

mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

#### TRANSPARENCY DECLARATION

The lead author (MG) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; no important aspects of the study have been omitted; and all discrepancies from the study as planned have been explained.

## **DATA SHARING STATEMENT**

All authors, external and internal, had full access to all of the data in the study and can take responsibility for the integrity of the data and the accuracy of the interpretation.

#### **DATA SHARING STATEMENT**

The data pertaining to this systematic review are available from the corresponding author upon reasonable request.

# **FIGURE CAPTIONS**

**Figure 1.** Flow of records through the selection process

#### **REFERENCES**

- Greig P, Snow R. Fatigue and risk: are train drivers safer than doctors? *BMJ* 2017;359. doi: https://doi.org/10.1136/bmj.j5107
- 2. Parker JBR. The effects of fatigue on physician performance an underestimated cause of physician impairment and increased patient risk. *Can J Anaesth* 1987;34(5):489-95.
- 3. Temple J. Resident duty hours around the globe: where are we now? BMC Med Educ 2014;14(Suppl 1):S8. https://doi.org/10.1186/1472-6920-14-S1-S8
- 4. Imrie K, Frank J, The National Steering Committee on Resient Duty Hours. Fatigue, risk, & excellence: towards a pan-Canadian consensus on resident duty hours. Ottawa, Ontario: The Royal College of Physicians and Surgeons of Canada, 2013. <a href="http://www.residentdutyhours.ca/final\_report.php">http://www.residentdutyhours.ca/final\_report.php</a> (accessed 12 Jan 2018).

- 6. Bolster L, Rourke L. The effect of restricting residents' duty hours on patient safety, resident well-being, and resident education: an updated systematic review. *J Grad Med Educ* 2015;7(3):349-63.
- 7. Harris JD, Staheli G, LeClere L, Andersone D, McCormick F. What effects have resident work-hour changes had on education, quality of life, and safety? A systematic review. *Clin Orthop* 2015;473(5):1600-8.
- 8. Ahmed N, Devitt KS, Keshet I, et al. A systematic review of the effects of resident duty hour restrictions in surgery: impact on resident wellness, training, and patient outcomes. *Ann Surg* 2014;259(6):1041-53.
- 9. Peets A, Ayas NT. Restricting resident work hours: the good, the bad, and the ugly. *Crit Care Med* 2012;40(3):960-6.
- Canadian Medical Association. CMA Policy: Management of physician fatigue. Ottawa, Canada:
   Canadian Medical Association; 2014.
- 11. Anim M, Markert RJ, Wood VC, Schuster BL. Physician practice patterns resemble ACGME duty hours. *Am J Med* 2009;122:587-93.
- 12. Transport Canada. Fatigue risk management system for the canadian aviation industry: policies and procedures development guidelines. Ottawa, Canada: Her Majesty the Queen in Right of Canada, as represented by the Minister of Transport, 2008.

  https://www.tc.gc.ca/eng/civilaviation/publications/TP14576-6042.htm (accessed 12 Jan 2018).
- 13. Patterson PD, Higgins JS, Lang ES, et al. Evidence-based guidelines for fatigue risk management in ems: formulating research questions and selecting outcomes. *Prehosp Emerg Care* 2017;21(2):149-56.
- 14. Boudreau RA, Grieco RL, Cahoon SL, Robertson RC, Wedel RJ. The pandemic from within: two surveys of physician burnout in Canada. *Can J Commun Ment Health* 2006;25(2):71-88.
- 15. Kumar S. Burnout and doctors: prevalence, prevention and intervention. *Healthcare* 2016;4(3):37. doi: 10.3390/healthcare4030037
- 16. Shanafelt TD, Boone S, Tan L, et al. Burnout and satisfaction with work-life balance among us physicians relative to the general us population. *Arch Intern Med* 2012;172(18):1377-85.
- 17. Brady KJS, Trockel MT, Khan CT, et al. What do we mean by physician wellness? A systematic review of its definition and measurement. *Acad Psychiatry* 2017. doi: 10.1007/s40596-017-0781-6

- 18. Fralick M, Flegel K. Physician burnout: who will protect us from ourselves? *CMAJ* 2014;186(10):731. doi: 10.1503/cmaj.140588
- 19. Shanafelt TD, Noseworthy JH. Executive leadership and physician well-being: nine organizational strategies to promote engagement and reduce burnout. *Mayo Clin Proc* 2017;92(1):129-46.
- 20. Lemaire JB, Wallace JE. Burnout among doctors. BMJ 2017;358:j3360. doi: 10.1136/bmj.j3360
- 21. Dyrbye LN, Trockel M, Frank E, et al. Development of a research agenda to identify evidence-based strategies to improve physician wellness and reduce burnout. *Ann Intern Med* 2017;166(10):743-4.
- 22. Higgins JPT, Green S (editors). The Cochrane handbook for systematic reviews of interventions, version 5.1.0. London, UK: The Cochrane Collaboration, 2011. <a href="http://www.handbook.cochrane.org">http://www.handbook.cochrane.org</a> (accessed 12 Jan 2018).
- 23. 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(7):e1000100. https://doi.org/10.1371/journal.pmed.1000100
- 24. Pattani R, Wu PE, Dhalla IA. Resident duty hours in Canada: past, present and future. *CMAJ* 2014;186(10):761-5.
- 25. International Monetary Fund. World economic outlook: too slow for too long. Washington, DC: International Monetary Fund, 2016. <a href="http://www.imf.org/external/pubs/ft/weo/2016/01/">http://www.imf.org/external/pubs/ft/weo/2016/01/</a> (accessed 12 Jan 2018).
- 26. Cochrane Effective Practice and Organisation of Care. Suggested risk of bias criteria for EPOC reviews. Oslo, Norway: Norwegian Knowledge Centre for the Health Services, 2016. http://epoc.cochrane.org/epoc-specific-resources-review-authors (accessed 12 Jan 2018).
- 27. Wells GA SB, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa, Canada: The Ottawa Hospital Research Institute, 2014. http://www.ohri.ca/programs/clinical\_epidemiology/oxford.asp (accessed 12 Jan 2018).
- 28. Amirian I, Andersen LT, Rosenberg J, Gogenur I. Laparoscopic skills and cognitive function are not affected in surgeons during a night shift. *J Surg Educ* 2014;71(4):543-50.
- 29. Aziz A. Sources of perceived stress among american medical doctors: a cross-cultural perspective. Cross Cultural Management 2004;11(4):28-39.
- 30. Beaujouan L, Czernichow S, Pourriat JL, Bonnet F. Prevalence and risk factors for substance abuse and dependence among anaesthetists: a national survey. *Ann Fr Anesth Réanim* 2005;24(5):471-9.

- 32. Chen I, Vorona R, Chiu R, Ware JC. A survey of subjective sleepiness and consequences in attending physicians. *Behav Sleep Med* 2008;6(1):1-15.
- 33. Chu MW, Stitt LW, Fox SA, Kiaii B, Quantz M, Guo L, et al. Prospective evaluation of consultant surgeon sleep deprivation and outcomes in more than 4000 consecutive cardiac surgical procedures. *Arch Surg* 2011;146(9):1080-5.
- 34. Doppia MA, Estryn-Béhar M, Fry C, Guetarni K, Lieutaud T, comité de pilotage de l'enquête SESMAT. Burnout in French doctors: a comparative study among anaesthesiologists and other specialists in French hospitals (SESMAT study). *Ann Fr Anesth Réanim* 2011;30(11):782-94.
- 35. Dutheil F, Trousselard M, Perrier C, et al. Urinary interleukin-8 is a biomarker of stress in emergency physicians, especially with advancing age--the JOBSTRESS randomized trial. *PLoS One* 2013;8(8):e71658. https://doi.org/10.1371/journal.pone.0071658
- 36. Ellman PI, Law MG, Tache-Leon C, et al. Sleep deprivation does not affect operative results in cardiac surgery. *Ann Thorac Surg* 2004;78(3):906-11.
- 37. Elovainio M, Heponiemi T, Jokela M, et al. Stressful work environment and wellbeing: what comes first? *J Occup Health Psychol* 2015;20(3):289-300.
- 38. Gander P, Millar M, Webster C, Merry A. Sleep loss and performance of anaesthesia trainees and specialists. *Chronobiol Int* 2008;25(6):1077-91.
- 39. Gander PH, Merry A, Millar MM, Weller J. Hours of work and fatigue-related error: a survey of New Zealand anaesthetists. *Anaesth Intensive Care* 2000;28(2):178-83.
- 40. 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(6):813-9.
- 41. Harbeck B, Suefke S, Haas CS, Lehnert H, Kropp P, Moenig H. No stress after 24-hour on-call shifts? *J Occup Health* 2015;57(5):438-47.
- 42. Heponiemi T, Puttonen S, Elovainio M. On-call work and physicians' well-being: testing the potential mediators. *Occup Med* 2014;64(5):352-7.
- 43. Jackson TN, Pearcy CP, Khorgami Z, Agrawal V, Taubman KE, Truitt MS. The physician attrition crisis: a cross-sectional survey of the risk factors for reduced job satisfaction among US surgeons. *World J Surg* 2017;24. doi: 10.1007/s00268-017-4286-y

- 44. Kaneita Y, Ohida T. Association of current work and sleep situations with excessive daytime sleepiness and medical incidents among Japanese physicians. *J Clin Sleep Med* 2011;7(5):512-22.
- 45. Lederer W, Kopp M, Hahn O, et al. Post-duty psychomotor performance in young and senior anaesthetists. *Eur J Anaesthesiol* 2006;23(3):251-6.
- 46. Leichtfried V, Putzer G, Perkhofer D, Schobersberger W, Benzer A. Circadian melatonin profiles during single 24-h shifts in anesthetists. *Sleep Breath* 2011;15(3):503-12.
- 47. Lindfors PM, Nurmi KE, Meretoja OA, et al. On-call stress among Finnish anaesthetists. *Anaesthesia* 2006;61(9):856-66.
- 48. Mahmood JI, Stoen Grotmol K, Tesli M, Vaglum P, Tyssen R. Contextual factors and mental distress as possible predictors of hazardous drinking in Norwegian medical doctors: a 15-year longitudinal, nationwide study. *Eur Addict Res* 2017;23(1):19-27.
- 49. Nishimura K, Nakamura F, Takegami M, et al. Cross-sectional survey of workload and burnout among Japanese physicians working in stroke care: the nationwide survey of acute stroke care capacity for proper designation of comprehensive stroke center in Japan (J-ASPECT) study. *Circ Cardiovasc Qual Outcomes* 2014;7(3):414-22.
- 50. Pit SW, Hansen V. Factors influencing early retirement intentions in Australian rural general practitioners. *Occup Med* 2014;64(4):297-304.
- 51. Pit SW, Hansen V. The relationship between lifestyle, occupational health, and work-related factors with presenteeism amongst general practitioners. *Arch Environ Occup Health* 2016;71(1):49-56.
- 52. Roberts DL, Shanafelt TD, Dyrbye LN, West CP. A national comparison of burnout and work-life balance among internal medicine hospitalists and outpatient general internists. *J Hosp Med* 2014;9(3):176-81.
- 53. Rothschild JM, Keohane CA, Rogers S, et al. Risks of complications by attending physicians after performing nighttime procedures. *JAMA* 2009;302(14):1565-72.
- 54. Saadat H, Bissonnette B, Tumin D, et al. Effects of partial sleep deprivation on reaction time in anesthesiologists. *Paediatr Anaesth* 2017;27(4):358-62.
- 55. Saadat H, Bissonnette B, Tumin D, et al. Time to talk about work-hour impact on anesthesiologists:

  The effects of sleep deprivation on Profile of Mood States and cognitive tasks. *Paediatr Anaesth*2016;26(1):66-71.
- 56. Sanches I, Teixeira F, dos Santos JM, Ferreira AJ. Effects of acute sleep deprivation resulting from night shift work on young doctors. *Acta Med Port* 2015;28(4):457-62.

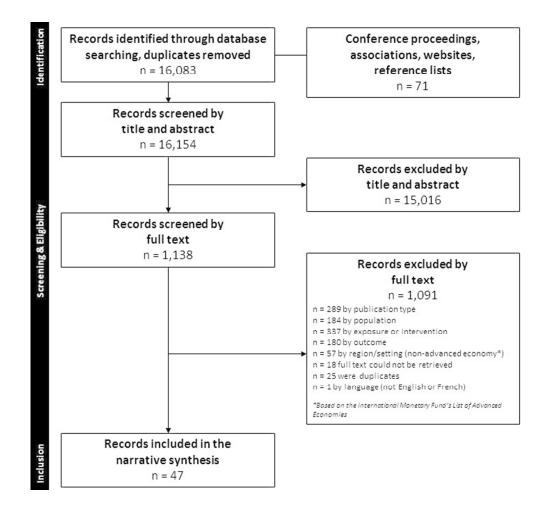
- 58. Schieman C, MacLean AR, Buie WD, Rudmik LR, Ghali WA, Dixon E. Does surgeon fatigue influence outcomes after anterior resection for rectal cancer? *Am J Surg* 2008;195(5):684-7.
- 59. Sende J, Jbeili C, Schvahn S, Khalid M, Asaph J, Romano H, et al. Stress factors and stress consequences among emergency physicians: vational survey. *Annales Francaises de Médecine d'Urgence* 2012;2(4):224-31.
- 60. Sexton JB, Thomas EJ, Helmreich RL. Error, stress, and teamwork in medicine and aviation: Cross sectional surveys. *Hum Perform Extrem Environ* 2001;6(1):6-11.
- 61. Shanafelt TD, Balch CM, Bechamps G, et al. Burnout and medical errors among American surgeons.

  Ann Surg 2010;251(6):995-1000.
- 62. Shanafelt TD, Novotny P, Johnson ME, et al. The well-being and personal wellness promotion strategies of medical oncologists in the North Central Cancer Treatment Group. *Oncology* 2005;68(1):23-32.
- 63. Shanafelt TD, Raymond M, Kosty M, et al. Satisfaction with work-life balance and the career and retirement plans of US oncologists. *J Clin Oncol* 2014;32(11):1127-35.
- 64. Shirom A, Nirel N, Vinokur AD. Overload, autonomy, and burnout as predictors of physicians' quality of care. *J Occup Health Psychol* 2006;11(4):328-42.
- 65. Shirom A, Nirel N, Vinokur AD. Work hours and caseload as predictors of physician burnout: the mediating effects by perceived workload and by autonomy. *J Appl Psychol* 2010;59(4):539-65.
- 66. Smith F, Goldacre MJ, Lambert TW. Adverse effects on health and wellbeing of working as a doctor: views of the UK medical graduates of 1974 and 1977 surveyed in 2014. *J R Soc Med* 2017;110(5):198-207.
- 67. Starmer AJ, Frintner MP, Freed GL. Work-life balance, burnout, and satisfaction of early career pediatricians. *Pediatrics* 2016;e20153183. doi: 10.1542/peds.2015-3183
- 68. Tanti A, Camilleri M, Borg AA, et al. Opinions of Maltese doctors and pharmacists on medication errors. *Int J Saf Med* 2017;29(1-2):81-99.
- 69. Tokuda Y, Hayano K, Ozaki M, Bito S, Yanai H, Koizumi S. The interrelationships between working conditions, job satisfaction, burnout and mental health among hospital physicians in Japan: a path analysis. *Ind Health* 2009;47(2):166-72.

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

- 70. 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(6):753-7.
- 71. Vela-Bueno A, Moreno-Jimenez B, Rodriguez-Munoz A, et al. Insomnia and sleep quality among primary care physicians with low and high burnout levels. *J Psychosom Res* 2008;64(4):435-42.
- 72. Vinden C, Nash DM, Rangrej J, et al. Complications of daytime elective laparoscopic cholecystectomies performed by surgeons who operated the night before. *Obstet Gynecol Surv* 2014;69(2):71-3.
- 73. Wada K, Yoshikawa T, Goto T, et al. National survey of the association of depressive symptoms with the number of off duty and on-call, and sleep hours among physicians working in Japanese hospitals: a cross sectional study. *BMC Public Health* 2010;10:127. doi: 10.1186/1471-2458-10-127
- 74. Govindarajan A, Urbach DR, Kumar M, et al. Outcomes of daytime procedures performed by attending surgeons after night work. *N Engl J Med* 2015;373(9):845-53.
- 75. Sturm L, Dawson D, Vaughan R, et al. Effects of fatigue on surgeon performance and surgical outcomes: a systematic review. *ANZ J Surg* 2011;81(7-8):502-9.
- 76. Wallace JE, Lemaire JB, Ghali WA. Physician wellness: a missing quality indicator. *Lancet* 2009;374(9702):1714-21.
- 77. Canadian Medical Association. CMA policy: physician health. Ottawa, Canada: Canadian Medical Assocation, 2017. <a href="http://policybase.cma.ca/dbtw-wpd/Policypdf/PD18-01.pdf">http://policybase.cma.ca/dbtw-wpd/Policypdf/PD18-01.pdf</a> (accessed 12 Jan 2018).
- 78. Shanafelt T, Goh J, Sinsky C. The business case for investing in physician well-being. *JAMA Intern Med* 2017;177(2):195-205.
- 79. Gaba DM, Howard SK. Fatigue among clinicians and the safety of patients. *N Engl J Med* 2002;347(16):1249-55.
- 80. Ward S, Outram S. Medicine: in need of culture change. Itern Med J 2016;46(1):112-6.
- 81. Panagioti M, Panagopoulou E, Bower P, et al. Controlled interventions to reduce burnout in physicians: a systematic review and meta-analysis. *JAMA Intern Med* 2017;177(2):195-205.
- 82. American Medical Association. STEPSforward<sup>TM</sup>. Chicago, IL: American Medical Assocation, 2017. <a href="https://www.stepsforward.org/modules/joy-in-medicine">https://www.stepsforward.org/modules/joy-in-medicine</a> (accessed 12 Jan 2018).
- 83. Jackson C. The Chalder Fatigue scale (CFQ 11). Occup Med 2015;65(1):86. doi: 10.1093/occmed/kgu168

- 85. Smets E, Garssen B, Bonke Bd, De Haes J. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *J Psychosom Res* 1995;39(3):315-25.
- 86. Ioannidis JPA, Greenland S, Hlatky MA, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;383(9912):166-75.
- 87. Gagliardi AR, Berta W, Kothari A, Boyko J, Urguhart R. Integrated knowledge translation (IKT) in ag review. Im. health care: a scoping review. Implement Sci 2016;11:38. https://doi.org/10.1186/s13012-016-0399-



60x56mm (300 x 300 DPI)

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text and

Date searched: 13 April 2016, updated 7 November 2017

Records retrieved: 5068 and 1442 in the update (removed duplicates retrieved in previous search)

- 1. Medical Staff, Hospital/
- 2. Physician Impairment/
- 3. exp Physicians/
- 4. allergist\*.ti.
- 5. (an?esthetist\* or an?esthesiologist\*).ti.
- 6. cardiologist\*.ti.
- 7. clinician\*.ti.
- 8. (clinician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 9. (clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw,kf.
- 10. dermatologist\*.ti.
- 11. endocrinologist\*.ti.
- 12. doctor\*.ti.
- 13. (doctor\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 14. (doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw,kf.
- 15. family practitioner\*.ti.
- 16. gastroenterologist\*.ti.
- 17. (general practitioner\* or GP\*).ti.
- 18. (general adj2 physician\*).ti.
- 19. geriatrician\*.ti.
- 20. gyn?ecologist\*.ti.

- 21. h?ematologist\*.ti.
- 22. (health\* adj2 (professional\* or provider\*)).ti.
- 23. hospitalist\*.ti.
- 24. (house staff\* or housestaff\*).ti.
- 25. intensivist\*.ti.
- 26. internist\*.ti.
- 27. medical professional\*.ti.
- 28. obstetrician\*.ti.
- 29. oncologist\*.ti.
- 30. ophthalmologist\*.ti.
- 31. orthop?edist\*.ti.
- 32. (otolaryngologist\* or otorhinolaryngologist\*).ti.
- 33. neonatologist\*.ti.
- 34. nephrologist\*.ti.
- 35. neurologist\*.ti.
- 36. neuropsychiatrist\*.ti.
- 37. neurosurgeon\*.ti.
- 38. p?ediatrician\*.ti.
- 39. perinatologist\*.ti.
- 40. physician\*.ti.
- 41. (physician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 42. (physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw,kf.
- 43. primary care practitioner\*.ti.
- 44. psychiatrist\*.ti.
- 45. pulmonologist\*.ti.
- 46. rheumatologist\*.ti.
- 47. surgeon\*.ti.

- 48. (surgeon\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 49. traumatologist\*.ti.
- 50. urologist\*.ti.
- 51. or/1-50 [Combined MeSH, title, and text word searches for physicians]
- 52. Burnout, Professional/
- 53. exp Circadian Rhythm/
- 54. exp Fatigue/
- 55. Occupational Health/
- 56. Rest/ph, px [Physiology, Psychology]
- 57. Sleep Deprivation/
- 58. Sleep Disorders, Circadian Rhythm/
- 59. Sleep Wake Disorders/
- 60. exp Stress, Psychological/
- 61. Workload/px [Psychology]
- 62. Work Schedule Tolerance/
- 63. ((24 hour\* or 24 hr\* or twenty four hour\* or twentyfour hour\*) adj rhythm\*).tw,kf.
- 64. biological rhythm\*.tw,kf.
- 65. (burn out\* or burned out\* or burnt out\* or burnout\*).tw,kf.
- 66. circadian misalignment.tw,kf.
- 67. ((circadian or diurnam or ultradian) adj rhythm\*).tw,kf.
- 68. exhaust\*.tw,kf.
- 69. fatigu\*.tw,kf.
- 70. (sleep\* adj3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)).tw,kf.
- 71. tired\*.tw,kf.
- 72. weariness.tw,kf.
- 73. or/52-72 [Combined MeSH and text words for fatigue]
- 74. and/51,73 [Combined concepts for physicians and fatigue]
- 75. animals/ not (animals/ and humans/)
- 76. 74 not 75
- 77. (comment or editorial or letter).pt.

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

- 78. 76 not 77
- 79. limit 78 to yr="2000-Current"
- 80. limit 79 to (english or french)
- 81. remove duplicates from 80

Database: Ovid Embase 1996 to 2016 Week 15

Date searched: 13 April 2016

Records retrieved: 8859

- 1. medical staff/
- 2. exp physician/
- 3. allergist\*.ti.
- 4. (an?esthetist\* or an?esthesiologist\*).ti.
- 5. cardiologist\*.ti.
- 6. clinician\*.ti.
- 7. (clinician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 8. (clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 9. dermatologist\*.ti.
- 10. endocrinologist\*.ti.
- 11. doctor\*.ti.
- 12. (doctor\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 13. (doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 14. family practitioner\*.ti.
- 15. gastroenterologist\*.ti.
- 16. (general practitioner\* or GP\*).ti.
- 17. (general adj2 physician\*).ti.

- 19. gyn?ecologist\*.ti.
- 20. h?ematologist\*.ti.
- 21. (health\* adj2 (professional\* or provider\*)).ti.
- 22. hospitalist\*.ti.
- 23. (house staff\* or housestaff\*).ti.
- 24. intensivist\*.ti.
- 25. internist\*.ti.
- 26. medical professional\*.ti.
- 27. obstetrician\*.ti.
- 28. oncologist\*.ti.
- 29. ophthalmologist\*.ti.
- 30. orthop?edist\*.ti.
- 31. (otolaryngologist\* or otorhinolaryngologist\*).ti.
- 32. neonatologist\*.ti.
- 33. nephrologist\*.ti.
- 34. neurologist\*.ti.
- 35. neuropsychiatrist\*.ti.
- 36. neurosurgeon\*.ti.
- 37. p?ediatrician\*.ti.
- 38. perinatologist\*.ti.
- 39. physician\*.ti.
- 40. (physician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 41. (physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 42. primary care practitioner\*.ti.
- 43. psychiatrist\*.ti.
- 44. pulmonologist\*.ti.
- 45. rheumatologist\*.ti.
- 46. surgeon\*.ti.

- 47. (surgeon\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 48. traumatologist\*.ti.
- 49. urologist\*.ti.
- 50. or/1-49 [Combined Emtree, title, and text word searches for physicians]
- 51. burnout/
- 52. circadian rhythm/
- 53. circadian rhythm sleep disorder/
- 54. fatigue/
- 55. mental stress/
- 56. occupational health/
- 57. sleep deprivation/
- 58. sleep waking cycle/
- 59. work capacity/
- 60. work schedule/
- 61. working time/
- 62. workload/
- 63. ((24 hour\* or 24 hr\* or twenty four hour\* or twentyfour hour\*) adj rhythm\*).tw.
- 64. biological rhythm\*.tw.
- 65. (burn out\* or burned out\* or burnt out\* or burnout\*).tw.
- 66. circadian misalignment.tw.
- 67. ((circadian or diurnam or ultradian) adj rhythm\*).tw.
- 68. exhaust\*.tw.
- 69. fatigu\*.tw.
- 70. (sleep\* adj3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)).tw.
- 71. tired\*.tw.
- 72. weariness.tw.
- 73. or/51-72 [Combined Emtree and text words for fatigue]
- 74. and/50,73 [Combined concepts for physicians and fatigue]
- 75. animals/ not (animals/ and humans/)
- 76. 74 not 75

77. (conference\* or editorial or letter or proceeding).pt.

- 78.76 not 77
- 79. limit 78 to yr="2000-Current"
- 80. limit 79 to (english or french)
- 81. limit 80 to embase

# Database: Ovid PsycINFO 1987 to April Week 1 2016

Date searched: 13 April 2016

Records retrieved: 2094

- 1. exp Physicians/
- 2. allergist\*.ti.
- 3. (an?esthetist\* or an?esthesiologist\*).ti.
- 4. cardiologist\*.ti.
- 5. clinician\*.ti.
- 6. (clinician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 7. (clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 8. dermatologist\*.ti.
- 9. endocrinologist\*.ti.
- 10. doctor\*.ti.
- 11. (doctor\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 12. (doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 13. family practitioner\*.ti.
- 14. gastroenterologist\*.ti.
- 15. (general practitioner\* or GP\*).ti.
- 16. (general adj2 physician\*).ti.

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

- 17. geriatrician\*.ti.
- 18. gyn?ecologist\*.ti.
- 19. h?ematologist\*.ti.
- 20. (health\* adj2 (professional\* or provider\*)).ti.
- 21. hospitalist\*.ti.
- 22. intensivist\*.ti.
- 23. internist\*.ti.
- 24. medical professional\*.ti.
- 25. obstetrician\*.ti.
- 26. oncologist\*.ti.
- 27. ophthalmologist\*.ti.
- 28. orthop?edist\*.ti.
- 29. (otolaryngologist\* or otorhinolaryngologist\*).ti.
- 30. neonatologist\*.ti.
- 31. nephrologist\*.ti.
- 32. neurologist\*.ti.
- 33. neuropsychiatrist\*.ti.
- 34. neurosurgeon\*.ti.
- 35. p?ediatrician\*.ti.
- 36. perinatologist\*.ti.
- 37. physician\*.ti.
- 38. (physician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 39. (physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 40. primary care practitioner\*.ti.
- 41. psychiatrist\*.ti.
- 42. pulmonologist\*.ti.
- 43. rheumatologist\*.ti.
- 44. surgeon\*.ti.

- 45. (surgeon\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 46. traumatologist\*.ti.
- 47. urologist\*.ti.
- 48. or/1-47 [Combined thesaurus, title, and text word searches for physicians]
- 49. Compassion Fatigue/
- 50. Fatigue/
- 51. Human Biological Rhythms/
- 52. Occupational Health/
- 53. Occupational Stress/
- 54. Sleep/
- 55. Sleepiness/
- 56. Working Conditions/
- 57. Work Rest Cycles/
- 58. Work Week Length/
- 59. Work Scheduling/
- 60. Workday Shifts/
- 61. ((24 hour\* or 24 hr\* or twenty four hour\* or twentyfour hour\*) adj rhythm\*).tw.
- 62. biological rhythm\*.tw.
- 63. (burn out\* or burned out\* or burnt out\* or burnout\*).tw.
- 64. circadian misalignment.tw.
- 65. ((circadian or diurnam or ultradian) adj rhythm\*).tw.
- 66. exhaust\*.tw.
- 67. fatigu\*.tw.
- 68. (sleep\* adj3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)).tw.
- 69. tired\*.tw.
- 70. weariness.tw.
- 71. or/49-70 [Combined thesaurus and text words for fatigue]
- 72. and/48,71 [Combined concepts for physicians and fatigue]
- 73. limit 72 to yr="2000-Current"
- 74. limit 73 to (english or french)

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

## Database: CINAHL Plus with Full Text (1937 to the present) via EBSCOhost

Date searched: 14 April 2016

Records retrieved: 3378

- S1. (MH "Medical Staff, Hospital+")
- S2. (MH "Physicians+")
- S3. TI allertist\*
- S4. TI (anesthetist\* or anaesthetist\* or anesthesiologist\* or anaesthesiologist\*)
- S5. TI cardiologist\*
- S6. TI clinician\*
- S7. clinician\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or "work\* hour\*" or "work life balance")
- S8. clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)
- S9. TI dermatologist\*
- S10. TI endocrinologist\*
- S11. TI doctor\*
- S12. doctor\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or work\* hour\* or "work life balance")
- S13. doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)
- S14. TI "family practitioner\*"
- S15. TI gastroenterologist\*
- S16. TI ("general practitioner\*" or GP\*)
- S17. TI (general N2 physician\*)
- S18. TI geriatrician\*
- S19. TI (gynaecologist\* or gynecologist\*)
- S20. TI (haematologist\* or hematologist\*)
- S21. TI hospitalist\*
- S22. TI ("house staff\*" or housestaff\*)

- S23. TI intensivist\*
- S24. TI internist\*

- S25. TI obstetrician\*
- S26. TI oncologist\*
- S27. TI ophthalmologist\*
- S28. TI (orthopaedist\* or orthopedist\*)
- S29. TI (otolaryngologist\* or otorhinolaryngologist\*)
- S30. TI neonatologist\*
- S31. TI nephrologist\*
- S32. TI neurologist\*
- S33. TI neuropsychiatrist\*
- S34. TI neurosurgeon\*
- S35. TI (paediatrician\* OR pediatrician\*)
- S36. TI perinatologist\*
- S37. TI physician\*
- S38. physician\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or "work\* hour\*" or "work life balance")
- S39. physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)
- S40. TI "primary care practitioner\*"
- S41. TI psychiatrist\*
- S42. TI pulmonologist\*
- S43. TI rheumatologist\*
- S44. TI surgeon\*
- S45. surgeon\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or work\* hour\* or "work life balance")
- S46. TI traumatologist\*
- S47. TI urologist\*
- S48. S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

```
OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41 OR
```

- S42 OR S43 OR S44 OR S45 OR S46 OR S47
- S49. (MH "Circadian Rhythm")
- S50. (MH "Fatigue")
- S51. (MH "Impairment, Health Professional")
- S52. (MH "Mental Fatigue")
- S53. (MH "Occupational Health")
- S54. (MH "Shiftwork")
- S55. (MH "Sleep Deprivation")
- S56. (MH "Sleep Disorders, Circadian Rhythm")
- S57. (MH "Sleep-Wake Transition Disorders")
- S58. (MH "Stress, Occupational+")
- S59. (MH "Stress, Psychological")
- S60. ("24 hour\*" or "24 hr\*" or "twenty four hour\*" or "twentyfour hour\*") N1 rhythm\*
- S61. "biological rhythm\*"
- S62. "burn out\*" or "burned out\*" or "burnt out\*" or burnout\*
- S63. "circadian misalignment"
- S64. (circadian or diurnam or ultradian) N1 rhythm\*
- S65. exhaust\*
- S66. fatigu\*
- S67. sleep\* N3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)
- S68. tired\*
- S69. weariness
- S70. S49 OR S50 OR S51 OR S52 OR S53 OR S54 OR S55 OR S56 OR S57 OR S58 OR S59 OR S60 OR S61 OR
- S62 OR S63 OR S64 OR S65 OR S66 OR S67 OR S68 OR S69
- S71. S48 AND S70
- S72. S48 AND S70 Limiters Published Date: 20000101-20161231; Publication Type: Clinical Trial, Journal
- Article, Meta Analysis, Meta Synthesis, Practice Guidelines, Randomized Controlled Trial, Research,

Review, Systematic Review; Language: English, French

**Database: PubMed via NCBI Entrez** 

Date searched: 14 April 2016

Records retrieved: 92

(((("Medical Staff, Hospital"[mh:noexp] OR "Physician Impairment"[mh:noexp] OR "Physicians"[mh] OR allergist[ti] OR allergists[ti] OR anaesthetist[ti] OR anaesthetists[ti] OR anaesthesiologist[ti] OR anaesthesiologists[ti] OR anesthetist[ti] OR anesthetists[ti] OR anesthesiologist[ti] OR anesthesiologists[ti] OR cardiologist[ti] OR cardiologists[ti] OR clinician[ti] OR clinicians[ti] OR ((clinician[tiab] OR clinicians[tiab]) AND (absent[tiab] OR absentee[tiab] OR absenteeism[tiab] OR absenteeisms[tiab] OR "burned out"[tiab] OR "burn out"[tiab] OR "burn outs"[tiab] OR burnout[tiab] OR burnouts[tiab] OR "burnt out"[tiab] OR coping[tiab] OR distress[tiab] OR distressed[tiab] OR distresses[tiab] OR distressing[tiab] OR "duty hour"[tiab] OR "duty hours"[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR health[tiab] OR healthy[tiab] OR impair[tiab] OR impaired[tiab] OR impairment[tiab] OR resilience[tiab] OR resiliency[tiab] OR resilient[tiab] OR satisfaction[tiab] OR sleep[tiab] OR sleepiness[tiab] OR sleeps[tiab] OR "well being"[tiab] OR wellbeing[tiab] OR wellness[tiab] OR "work hours"[tiab] OR "working hours"[tiab] OR "work life balance"[tiab])) OR ((clinician[tiab] OR clinicians[tiab]) AND (cities[tiab] OR city[tiab] OR community[tiab] OR communities[tiab] OR country[tiab] OR countryside[tiab] OR frontier[tiab] OR north[tiab] OR northern[tiab] OR remote[tiab] OR rural[tiab] OR suburb[tiab] OR suburbs[tiab] OR suburban[tiab] OR urban[tiab] OR urbanite[tiab])) OR dermatologist[ti] OR dermatologists[ti] OR endocrinologist[ti] OR endocrinologists[ti] OR doctor[ti] OR doctors[ti] OR ((doctor[tiab] OR doctors[tiab]) AND (absent[tiab] OR absentee[tiab] OR absenteeism[tiab] OR absenteeisms[tiab] OR "burned out"[tiab] OR "burn out"[tiab] OR "burn outs"[tiab] OR burnout[tiab] OR burnouts[tiab] OR "burnt out"[tiab] OR coping[tiab] OR distress[tiab] OR distressed[tiab] OR distresses[tiab] OR distressing[tiab] OR "duty hour"[tiab] OR "duty hours"[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR health[tiab] OR healthy[tiab] OR impair[tiab] OR impaired[tiab] OR impairment[tiab] OR resilience[tiab] OR resiliency[tiab] OR resilient[tiab] OR satisfaction[tiab] OR sleep[tiab] OR sleepiness[tiab] OR sleeps[tiab] OR "well being"[tiab] OR wellbeing[tiab] OR wellness[tiab] OR "work hours"[tiab] OR "working hours"[tiab] OR "work life balance"[tiab])) OR ((doctor[tiab] OR doctors[tiab]) AND (cities[tiab] OR city[tiab] OR community[tiab] OR communities[tiab] OR country[tiab] OR countryside[tiab] OR frontier[tiab] OR north[tiab] OR northern[tiab] OR remote[tiab] OR rural[tiab] OR suburb[tiab] OR suburbs[tiab] OR suburban[tiab] OR urban[tiab] OR urbanite[tiab])) OR "family practitioner"[ti] OR

"family practitioners"[ti] OR gastroenterologist[ti] OR gastroenterologists[ti] OR "general practice physician"[ti] OR "general practice physicians"[ti] OR "general practitioner"[ti] OR "general practitioners"[ti] OR geriatrician[ti] OR geriatricians[ti] OR gynaecologist[ti] OR gynaecologists[ti] OR gynecologist[ti] OR gynecologists[ti] OR haematologist[ti] OR haematologists[ti] OR hematologist[ti] OR hematologists[ti] OR "health care professional"[ti] OR "health care professionals"[ti] AND "health care provider"[ti] OR "health care providers" OR "health professional"[ti] OR "health professionals"[ti] OR "health provider"[ti] OR "health providers"[ti] OR "healthcare professional"[ti] OR "healthcare professionals"[ti] OR "healthcare provider"[ti] OR "healthcare providers"[ti] OR hospitalist[ti] OR hospitalists[ti] OR "house staff"[ti] OR "house staffs"[ti] OR housestaff[ti] OR housestaffs[ti] OR intensivist[ti] OR intensivists[ti] OR internist[ti] OR internists[ti] OR "medical professional"[ti] OR "medical professionals"[ti] OR obstetrician[ti] OR obstetricians[ti] OR oncologist[ti] OR oncologists[ti] OR ophthalmologist[ti] OR ophthalmologists[ti] OR orthopaedist[ti] OR orthopaedists[ti] OR orthopaedist[ti] OR orthopedists[ti] OR otolaryngologist[ti] OR otolaryngologists[ti] OR otorhinolaryngologist[ti] OR otorhinolaryngologists[ti] OR neonatologist[ti] OR neonatologists[ti] OR nephrologist[ti] OR nephrologists[ti] OR neurologist[ti] OR neurologists[ti] OR neuropsychiatrist[ti] OR neuropsychiatrists[ti] OR neurosurgeon[ti] OR neurosurgeons[ti] OR paediatrician[ti] OR paediatricians[ti] OR pediatrician[ti] OR pediatricians[ti] OR perinatologist[ti] OR perinatologists[ti] OR physician[ti] OR physicians[ti] OR ((physician[tiab] OR physicians[tiab]) AND (absent[tiab] OR absentee[tiab] OR absenteeism[tiab] OR absenteeisms[tiab] OR "burned out"[tiab] OR "burn out"[tiab] OR "burn outs"[tiab] OR burnout[tiab] OR burnouts[tiab] OR "burnt out"[tiab] OR coping[tiab] OR distress[tiab] OR distressed[tiab] OR distresses[tiab] OR distressing[tiab] OR "duty hour"[tiab] OR "duty hours"[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR health[tiab] OR healthy[tiab] OR impair[tiab] OR impaired[tiab] OR impairment[tiab] OR resilience[tiab] OR resiliency[tiab] OR resilient[tiab] OR satisfaction[tiab] OR sleep[tiab] OR sleepiness[tiab] OR sleeps[tiab] OR "well being"[tiab] OR wellbeing[tiab] OR wellness[tiab] OR "work hours"[tiab] OR "working hours"[tiab] OR "work life balance"[tiab])) OR ((physician[tiab] OR physicians[tiab]) AND (cities[tiab] OR city[tiab] OR community[tiab] OR communities[tiab] OR country[tiab] OR countryside[tiab] OR frontier[tiab] OR north[tiab] OR northern[tiab] OR remote[tiab] OR rural[tiab] OR suburb[tiab] OR suburbs[tiab] OR suburban[tiab] OR urban[tiab] OR urbanite[tiab])) OR "primary care practitioner"[ti] OR "primary care practitioners"[ti] OR psychiatrist[ti] OR psychiatrists[ti] OR pulmonologist[ti] OR pulmonologists[ti] OR rheumatologist[ti] OR rheumatologists[ti] OR surgeon[ti] OR surgeons[ti] OR ((surgeon[tiab] OR surgeons[tiab]) AND (absent[tiab] OR absentee[tiab] OR absenteeism[tiab] OR absenteeisms[tiab] OR

"burned out"[tiab] OR "burn out"[tiab] OR "burn outs"[tiab] OR burnout[tiab] OR burnouts[tiab] OR "burnt out"[tiab] OR coping[tiab] OR distress[tiab] OR distressed[tiab] OR distresses[tiab] OR distressing[tiab] OR "duty hour"[tiab] OR "duty hours"[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR health[tiab] OR healthy[tiab] OR impair[tiab] OR impaired[tiab] OR impairment[tiab] OR resilience[tiab] OR resiliency[tiab] OR resilient[tiab] OR satisfaction[tiab] OR sleep[tiab] OR sleepiness[tiab] OR sleeps[tiab] OR "well being"[tiab] OR wellbeing[tiab] OR wellness[tiab] OR "work hours"[tiab] OR "working hours"[tiab] OR "work life balance"[tiab])) OR traumatologist[ti] OR traumatologists[ti] OR urologist[ti] OR urologists[ti]) AND ("Burnout, Professional"[mh:noexp] OR "Circadian Rhythm"[mh] OR "Fatigue"[mh] OR "Occupational Health"[mh:noexp] OR "Rest/physiology"[mh:noexp] OR "Rest/psychology"[mh:noexp] OR "Sleep Deprivation"[mh:noexp] OR "Sleep Disorders, Circadian Rhythm"[mh:noexp] OR "Stress, Psychological"[mh] OR "Workload/psychology"[mh] OR "Work Schedule Tolerance"[mh:noexp] OR "24 hour rhythm"[tiab] OR "24 hour rhythms"[tiab] OR "24 hr rhythm"[tiab] OR "24 hr rhythms"[tiab] OR alertness[tiab] OR "biological rhythm"[tiab] OR "biological rhythms"[tiab] OR "burn out"[tiab] OR "burned out"[tiab] OR "burnt out"[tiab] OR burnout[tiab] OR "circadian misalignment"[tiab] OR "circadian rhythm"[tiab] OR "circadian rhythms"[tiab] OR "diurnal rhythm"[tiab] OR "diurnal rhythms"[tiab] OR exhausted[tiab] OR exhaustion[tiab] OR exhausting[tiab] OR exhausts[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR (("Sleep"[mh:noexp] OR sleep[tiab] OR sleeping[tiab]) AND (deprivation[tiab] OR deprive[tiab] OR deprived[tiab] OR deprives[tiab] OR depriving[tiab] OR disorder[tiab] OR disorders[tiab] OR lack[tiab] OR lacked[tiab] OR lacking[tiab] OR lacks[tiab] OR loss[tiab] AND insufficient[tiab] OR problem[tiab] OR problems[tiab])) OR tired[tiab] OR tiredness[tiab] OR "twenty four hour rhythm"[tiab] OR "twenty four hour rhythms"[tiab] OR weariness[tiab] OR "ultradian rhythm"[tiab] OR "ultradian rhythms"[tiab])) NOT (((Animals[MESH] OR Animal Experimentation[MESH] OR "Models, Animal"[MESH] OR Vertebrates[MESH]) NOT (Humans[MESH] OR Human experimentation[MESH])) OR (((animals[tiab] OR animal model[tiab] OR rat[tiab] OR rats[tiab] OR mouse[tiab] OR mice[tiab] OR rabbit[tiab] OR rabbits[tiab] OR pig[tiab] OR pigs[tiab] OR porcine[tiab] OR swine[tiab] OR dog[tiab] OR dogs[tiab] OR hamster[tiab] OR hamsters[tiab] OR chicken[tiab] OR chickens[tiab] OR sheep[tiab]) AND (publisher[sb] OR inprocess[sb] OR pubmednotmedline[sb])) NOT (human[ti] OR humans[ti] OR people[ti] OR children[ti] OR adults[ti] OR seniors[ti] OR patient[ti] OR patients[ti])))) NOT (editorial[pt] OR comment[pt] OR letter[pt] OR newspaper article[pt])) AND ((publisher[sb] NOT pubstatusnihms NOT pubstatuspmcsd NOT pmcbook) OR (pubstatUSheadofprint))

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

## **Supplementary table 1.** Descriptive characteristics of the included studies

Study	Physician and patient chara	cteristics			Setting		Interventions or	Outcomes	
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures		
Observational (exp	posure) studies (n=45)								
Cohort design									
Chu, 2011 [33]	Surgeons	6	NR	Range: 32-55y	Tertiary care academic	Urban	Sleep deprivation due	Surgeon operative	
Canada	Patients: cardiac surgery cases	4,047	NR	NR	¯ hospital		to work on the night preceding surgery	efficiency; patient postoperative mortality, adverse outcomes, length of stay	
Ellman, 2004 [36]	Surgeons	NR	NR	NR	University hospitals	Urban	Sleep deprivation due	Surgeon operative	
US	Patients: adult cardiac surgery cases	6,751	70%	S: 63.4±0.7y C: 63.5±0.1y	-		to work on the night preceding surgery	efficiency; patient complications, in-hospital mortality, length of stay, need for blood products	
Govindarajan,	Surgeons	1,448	NR	46.3±8.7	Academic and non-	Mixed	Sleep deprivation due	Duration of surgery;	
2015 [74] Canada	Patients: surgical cases	38,978	NR	56.4±16.6y	academic hospitals		to work on the night preceding a daytime surgery	Patient complications, mortality, readmissions, length of stay	
Rothschild, 2009	Surgeons	220	Surgeons:	Surgeons:	Tertiary care academic	Urban	Sleep deprivation due	Patient complications,	
[53] US	Obstetrician/gynecologists		84% OB/GYNs: 28%	42.0±7.6y OB/GYNs: 42.0±9.0y	trauma centre/referral centre for high-risk obstetrics		to work on the night preceding a daytime procedure	preventable complications	
	Patients: surgical and	Surg.:	Surg:	Surg:	_				
	obstetrics cases	4,471	S: 25%	S: 49.1±16.3y					
		Obst.:	C: 28%	C: 50.0±16.3y					
		4,902	Obst.:	Obst.:					
			S: 0%	S: 32.9±5.2y					
			C: 0%	C: 33.5±5.0y					
Schieman, 2007 [58]	Colorectal surgeons	NR	NR	NR	University teaching hospitals	NR	Fatigue due to work on the night preceding	Patient operative complications, length of	
Canada	Patients: undergoing anterior resection for rectal cancer	270	NR	S: 64.5y C: 64.4y	_		surgery	stay, mortality, cancer recurrence	

Study	Physician and patient chara	cteristics			Setting		Interventions or	Outcomes	
Country	Туре	n=	Sex (% male)	Age	Location	Urban	exposures		
						or			
						rural			
Vinden, 2014 [72]	General surgeons	331	83%	48±10y	Community hospitals	Mixed	Sleep deprivation due	Patient mortality,	
Canada	Patients: Elective	10,390	S: 27%	S: 49±16y			to overnight work	operative complications	
	cholecystectomies		C: 26%	C: 49±16y			preceding daytime		
							surgery		
Before-after design									
Amirian, 2014 [28]	Surgeons	29	55%	Median: 35y	Academic hospital	Urban	17-h night shift with	Cognitive and	
Denmark				Range: 27-49y			sleep deprivation	psychomotor abilities on	
			<b>/</b>					laparoscopic simulation	
Gerdes, 2008 [40]	Surgeons	9	NR	NR	University Hospital	Urban	Fatigue; sleep	Cognitive and	
US							deprivation from	psychomotor abilities	
							overnight call shift		
Lederer, 2006 [45]	Senior anesthetists	11	82%	49.0±2.0y	Hospital	Urban	Sleep deprivation from	Concentration ability;	
Austria							24-h call shift	reaction time;	
								performance on	
								psychometric tasks	
Time series design									
Leichtfried, 2011	Anesthetists	10	100%	Mean: 32y	<b>University Hospital</b>	Urban	Sleep deprivation from	Melatonin metabolite	
[46]				Range: 29-35y			24-h shift; sleepiness,	profile	
Austria							sleep hours		
Cross-sectional des	<u> </u>								
Aziz, 2004 [29]	Family medicine physicians	153	NR	NR	Hospitals	NR	Fatigue	Stress	
US	Various specialties								
Beaujouan, 2005	Anesthesiologists	3,476	64%	≤35y: 9%	Public sector	NR	Sleep deprivation	Substance abuse	
[30]				36-45y: 28%	General hospitals				
France				46-55y: 49%	University hospitals				
				56-65y: 13%	Private hospitals				
Chang, 2013 [31]	Anesthesiologists	11	64%	Mean: 38y	Level 1 trauma centre	NR	Sleep deprivation due	Cognitive performance;	
US			3.70	IQR: 34-48y	20.0. I dadina centre		to 15-h overnight call	reaction time	
				14,11, 54 40)			shift; sleepiness	reaction time	
Chen, 2008 [32]	Psychiatrists	180	77%	Academic:	Medical school	Urban	Sleep deprivation;	Impact on personal and	
US	Internists			79% 36-55v	Private practices		sleepiness	professional life;	

Study	Physician and patient chara	cteristics			Setting		Interventions or	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures	
	General practitioners Surgeons Obstetrician-gynecologists Radiologists Pediatricians Other			Private practice: 73% 36-65y				perceived risk of errors
Doppia, 2011 [34] France	Anesthesiologists	565	64%	<35y: 11% 35-54y: 63% >55y: 25%	Public hospitals Private hospitals Work-health environments Public health units	NR	Sleep deprivation	Burnout
Elovaino, 2015 [37] Finland	Physicians in various specialties	1,524	40%	Median: 49.7y Range: 24-69y	Hospitals Primary care Private practice Other unspecified	NR	Sleep difficulties	Job demands and control
Gander, 2000 [39] New Zealand	Anesthetists	183	NR	Mean: 46y	Combined public/private practice Other unspecified	NR	Work-related sleep disturbance	Risk of fatigue-related errors
Harbeck, 2015 [41] Germany	Internists	20	45%	Median: 32y Range: 26-42y	Hospital	NR	Sleep disturbance due to a 24-call shift	Biochemical and physiological parameters neurocognitive function
Heponiemi, 2014 [42] Finland	Physicians in various specialties Non-specialized physicians	1,541	40%	49.80±9.49y, Range: 24-67y	Hospitals Primary care clinic Private practice Other unspecified	NR	Sleep difficulties	Job satisfaction; work ability; psychological distress
Jackson, 2017 [43] US	Surgeons in various subspecialties	993	61%	More; less satisfied: 30-39y: 23%;24% 40-49y: 32%;36% 50-59y:	Academic practice Non-academic practice	NR	Not feeling well rested	Job satisfaction

Study	Physician and patient chara			Setting		Interventions or	Outcomes	
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures	
				23%;27% ≥60y: 23%;14%				
Kanieta, 2011 [44] Japan	Internists Surgeons Orthopedics Pediatricians Obstetrician-gynecologists Psychiatrists Dermatologists Urologists Opthalmologists Otorhinolaryngologists Other	3,486	66%	20-39y: 11% 40-49y: 25% 50-59y: 28% 60-69y: 16% ≥70y: 21%	Hospitals Clinics Other unspecified	NR	Sleep deprivation and difficulties; insomnia	Medical incidents
Lindfors, 2006 [47] Finland	Anesthetists	328	53%	47±7.8y Range: 32-69y	University hospitals Central and district hospitals Private sector	NR	Sleep disturbances; sleepiness	Stress; suicidal tendencies
Mahmood, 2016 [48] Norway	Generalists Internists Pediatricians Surgical specialties Anesthesiologists	450 (all time points)	41%	43y±2.8y	Public health system Private practice	NR	Sleep deprivation due to on-call shifts	Alcohol misuse
Nishimura, 2014 [49] Japan	Neurosurgeons and neurologists	2,564	NR	NR	Stroke care centres Teaching hospitals	NR	Sleep deprivation	Burnout
Pit, 2014 [50] Australia	General practitioners	92	60%	50±10.7y	NR	Rural	Work-related sleep disturbance	Early retirement intentions
Pit, 2016 [51] Australia	General practitioners	92	60%	50±10.7y	Private (solo) practice Group practice	Rural	Work-related sleep disturbance	Sickness presenteeism
Roberts, 2014 [52] US	General internists Internal medicine	578	58%	Hospitalists: 46.9±12.4y	Private practice Academic medical	NR	Fatigue	Falling asleep while driving

Study	Physician and patient cha	racteristics			Setting		Interventions or	Outcomes	
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures		
	hospitalists			Generalists: 53.6±10.2y	centre Veterans hospital Military practice Other				
Saadat, 2016 [55] US	Anesthesiologists	21	71%	30-40y: 57% 41-50y: 19% 51-55y: 24% Range: 32-56y	Tertiary care academic children's hospital	NR	Sleep deprivation due to 17-h night call shift	Mood disturbances	
Saadat, 2017 [54] US	Anesthesiologists	21	65%	Range: 32-56 years	Tertiary care academic children's hospital	NR	Sleep deprivation due to 17-h night call shift	Reaction time	
Sanches, 2015 [56] Spain	Emergency medicine physicians	18	28%	29.2±2.6y	Central hospital	NR	Sleep deprivation	Cognitive and psychomotor abilities	
Sargent, 2009 [57] US	Orthopedic surgeons	264	92%	NR	Orthopedic surgery training programs	NR	Sleep deprivation	Burnout; psychological distress; marital satisfaction	
Sende, 2012 [59] France	Emergency physicians	318	62%	39±8y	Hospitals Mobile emergency services Other unspecified	NR	Fatigue; sleep deprivation	Stress	
Sexton, 2001 [60] US	Consulting physicians: Surgeons Anesthesiologists Pulmonary physicians Cardiologists Pediatricians	271	NR	NR	Teaching and non- teaching hospitals	Urban	Fatigue	Perceived performance effectiveness	
Shanafelt, 2005 [62] US, Canada, Mexico	Oncologists	241	85%	>50y: 51%	Community clinics Hospitals Private practice Academic medical centres	NR	Fatigue; sleep deprivation	Quality of life/well-being	
Shanafelt, 2010	Surgeons	7,905	87%	Median: 51y	Private practice	NR	Fatigue	Perceived major medica	

Study	Physician and patient cha	racteristics			Setting		Interventions or	Outcomes	
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures		
[61] US				Q1: 43y Q2: 59y	Academic medical centres Veterans hospital Active military practice			errors	
					Retired or not in practice Other				
Shanafelt, 2014 [63] US	Oncologists	1,117	52%	Median: 52y	Private practice Academic practice Veteran's hospital Industry, other	NR	Fatigue	Satisfaction with work-life balance	
Shirom, 2006 [64] Israel	Opthalmologists Dermatologists Otolaryngologists Gynecologists General surgeons Cardiologists	890	80%	Median: 52y SD: 7.2y	Community clinics Acute care hospital outpatient clinics	NR	Physical fatigue	Perception of quality of patient care	
Shirom, 2010 [65] Israel	Opthalmologists Dermatologists Otolaryngologists Gynecologists General surgeons Cardiologists	890	80%	Median: 52y SD: 7.2y	Community clinics Acute care hospital outpatient clinics	NR	Physical fatigue	Burnout	
Smith, 2017 [66] UK	General practitioners Surgeons Other unspecified specialties	3,550	63%	NR	NR (varied)	NR	Perceived fatigue, sleep deprivation	Physical and mental health; competence	
Starmer, 2016 [67] US	General pediatricians Pediatric surgeons Pediatric hospitalists Pediatric specialists (unspecified)	840	40%	NR	NR (some in private practice)	NR	Sleep deprivation	Burnout; balanced personal and professional commitments; life and career satisfaction	
Tanti, 2017 [68]	Physicians (unspecified)	204	62%	Median: 41y	Hospitals	NR	Fatigue	Prescribing errors	

Study	Physician and patient char	acteristics			Setting		Interventions or	Outcomes	
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures		
Malta					Community Office-based				
Tokuda, 2009 [69] Japan	Hospital physicians: Generalists Other unspecified specialties	236	75%	40.9±7.8y Range: 26-76y	Hospitals with ≥20 inpatient beds	NR	Sleep deprivation	Burnout; job satisfaction	
Vela-Bueno, 2008 [71] Spain	Primary care physicians	113	27%	41.4±8.0y	Primary care centres	Urban	Sleep problems, insomnia	Burnout	
Wada, 2010 [73] Japan	Physicians (unspecified)	3,862	78%	M: 75% 30- 59y F: 85% 30-59y	Hospitals	NR	Sleep deprivation	Depressive symptoms	
Non-comparative of	design								
Gander, 2008 [38] New Zealand	Anesthetists	20	85%	Median: 44y	Hospitals	Urban	Sleep disturbance from consecutive working days or on- call work	Psychomotor performance	
Intervention studie	es (n=2)								
Randomized contro	olled trials								
Dutheil, 2013 [35] France	Emergency physicians	17	35%	39.1y±6.9y	University hospital	Urban	Fatigue related to 14-h and 24-h shifts; sleep deprivation; low sleep quality;	Perceived stress; urine interleukine-8	
Uchal, 2005 [70] Norway	Surgeons Gynecologists Orthopedic surgeons Urologists Vascular surgeons	64	67%	Median: Post-call: 33.0y Post-work: 38.0y	Government hospitals	NR	Sleep deprivation due to 24-h call shift	Product quality, procedure effectiveness of a surgical simulation	

C: control group; F: female; h: hour(s); IQR: interquartile range; M: male; NR: not reported; S: study group; SD: standard deviation; Surg: surgical; Obst: obstetric; Q: quartile; UK: United Kingdom; US: United States of America; y: year(s)

Supplementary table 2. Summary of risk of bias assessments for randomized controlled trials (n=2)<sup>a</sup>

First Author, Year	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other sources of bias	Overall risk of bias <sup>b</sup>
Dutheil, 2013	Low	Unclear	High	High	Low	Low	High	High
Uchal, 2005	Low	Low	Unclear	Low	Low	Low	Low	Unclear

<sup>&</sup>lt;sup>a</sup>Assessed using the Cochrane Collaboration's Risk of Bias Tool

Supplementary table 3. Summary of quality assessments for cohort studies (n=6)<sup>a</sup>

First Author,	Selection					Comparab	ility		Outco	me		Total
Year	Representa- tiveness of exposed cohort /1	Selection of non- exposed cohort /1	Ascertain- ment of exposure /1	Outcome not present at start /1	Total /4	Compara -bility of cohorts /2	Total /2	Assess- ment of outcome /1	Adequate length of follow-up /1	Adequate follow-up of cohorts /1	Total /1	Score <sup>b</sup> /9
Chu, 2011	1	1	0	1	3	2	2	1	1	1	3	8
Ellman, 2004	1	1	1	1	4	1	1	1	1	1	3	8
Govindarajan, 2015	1	1	1	1	4	2	2	1)	1	1	3	9
Rothschild, 2009	1	1	1	1	4	2	2	1	1	1	3	9
Schieman, 2008	1	1	1	1	4	1	1	1	1	1	3	8
Vinden, 2014	1	1	1	1	4	1	1	1	1	1	3	8

<sup>&</sup>lt;sup>a</sup>Assessed using the Newcastle-Ottawa Quality Assessment Scale

<sup>&</sup>lt;sup>b</sup>Overall risk of bias is Low if all domains are rated as low, High if at least one domain is assessed as high, and Unclear if at least one domain is assessed as unclear and no domains are assessed as high

<sup>&</sup>lt;sup>b</sup>An overall score of 7 to 9 stars is considered as low risk of bias, 4 to 6 as unclear risk of bias, and 3 or less as high risk of bias

## Supplementary table 4. Summary of risk of bias assessments for before-after studies (n=3)<sup>a</sup>

First Author, Year	Random sequence generation <sup>b</sup>	Allocation concealment <sup>b</sup>	Blinding of participants and	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other sources of bias <sup>c</sup>	Overall risk of bias <sup>d</sup>
			personnel					
Amirian, 2014	NA	NA	High	High	Low	Low	High	High
Gerdes, 2008	NA	NA	High	High	Low	Low	High	High
Lederer, 2006	NA	NA	High	High	Low	Low	High	High

<sup>&</sup>lt;sup>a</sup>Assessed using Cochrane Effective Practice and Organization of Care (EPOC) Review Group's criteria for before-after studies, adapted from the Cochrane Collaboration Risk of Bias Tool

## Supplementary table 5. Summary of risk of bias assessments for time series studies (n=1)<sup>a</sup>

	First Author, Year	Intervention independent of other changes	Intervention effect pre- specified	Intervention unlikely to affect data collection	Allocation concealment <sup>a</sup>	Incomplete outcome data	Selective reporting	Other sources of bias <sup>c</sup>	Overall risk of bias <sup>d</sup>
-	Leitchfried, 2011	Low	High	Low	NA	Low	Low	High	High

<sup>&</sup>lt;sup>a</sup>Assessed using Cochrane Effective Practice and Organization of Care (EPOC) Review Group's criteria for interrupted time series studies, adapted from the Cochrane Collaboration Risk of Bias Tool

<sup>&</sup>lt;sup>b</sup>Assessed as 'not applicable' (NA) when the studies did not include a control group

<sup>&</sup>lt;sup>c</sup>Assessed as High due to lack of a control group

Overall risk of bias is Low if all domains are rated as low, High if at least one domain is assessed as high, and Unclear if at least one domain is assessed as unclear and no domains are assessed as high

<sup>&</sup>lt;sup>b</sup>Assessed as not applicable (NA) when the studies did not include a control group

<sup>&</sup>lt;sup>c</sup>Assessed as High due to lack of a control group

Overall risk of bias is Low if all domains are rated as low, High if at least one domain is assessed as high, and Unclear if at least one domain is assessed as unclear and no domains are assessed as high

**Supplementary table 6.** Summary of quality assessments for cross-sectional studies (n=34)<sup>a</sup>

First Author, Year		Selection			Outcome			Total Score <sup>b</sup>
	Adequacy of case definition	Representative- ness of the sample /1	Total /2	Assessment of outcome /1	Same method of ascertainment for entire sample /1	Response rate /1	Total /3	/5
Aziz, 2004	0	0	0	0	1	0	1	1
Beaujouan, 2005	1	0	1	0	1	0	1	2
Chang, 2013	1	0	1	0	1	1	2	3
Chen, 2008	1	0	1	0	1	0	1	2
Doppia, 2011	1	1	2	0	1	1	2	4
Elovaino, 2015	1	1	2	0	1	1	2	4
Gander, 2000	1	1	2	0	1	1	2	4
Harbeck, 2015	1	0	1	0	1	1	2	3
Heponiemi, 2014	1	1	2	0	1	1	2	4
Jackson, 2017	0	0	0	0	1	0	1	1
Kanieta, 2011	1	0	1	0	1	1	2	3
Lindfors, 2006	1	1	2	0	1	1	2	4
Mahmood, 2017	1	0	1	0	1	0	1	2
Nishimura, 2014	1	1	2	0	1	0	1	3
Pit, 2014	1	0	1	0	1	1	2	3
Pit, 2016	1	0	1	0	1	1	2	3
Roberts, 2014	1	1	2	0	1	0	1	3
Saadat, 2016	1	1	2	0	1	1	2	4
Saadat, 2017	1	1	2	0	1	1	2	4
Sanches, 2015	1	0	1	0	1	0	2	3
Sargent, 2009	1	0	1	0	1	0	1	2

First Author, Year		Selection		Outcome				Total Score <sup>b</sup>
	Adequacy of case definition	Representative- ness of the sample /1	Total /2	Assessment of outcome /1	Same method of ascertainment for entire sample /1	Response rate /1	Total /3	/5
Sende, 2010	1	0	1	0	1	0	1	2
Sexton, 2001	1	0	1	0	1	0	1	2
Shanafelt, 2005	1	0	1	0	1	1	2	3
Shanafelt, 2010	1	0	1	0	1	0	1	2
Shanafelt, 2014	1	0	1	0	1	1	2	3
Shirom, 2006	1	1	2	0	1	1	2	4
Shirom, 2010	1	1	2	0	1	1	2	4
Smith, 2016	1	0	1	0	1	1	2	3
Starmer, 2016	1	1	2	0	1	1	2	4
Tanti, 2017	1	0	1	0	1	0	1	2
Tokuda, 2009	1	1	2	0	1	1	2	4
Vela-Bueno, 2008	1	1	2	0	1	1	2	4
Wada, 2010	1	1	2	0	1	0	1	3

<sup>&</sup>lt;sup>a</sup>Assessed using the Newcastle-Ottawa Quality Assessment Scale, adapted for cross-sectional studies

<sup>&</sup>lt;sup>b</sup>An overall score of 4 to 5 stars is considered as low risk of bias, 3 as unclear risk of bias, and 2 or less as high risk of bias. For response rate, ≥50% was used as the criterion to be awarded a star

Supplementary table 7. Summary of quality assessments for non-comparative studies (n=1)<sup>a</sup>

First Author,	, Selection			Exposu	Exposure			Outcome		
Year	Adequacy	Representat-	Total	Ascertain-	Total	Assessment	Same method	Loss to	Total	Score <sup>b</sup>
	of case	iveness of	/2	ment of	/1	of outcome	of assessment	follow-up	/3	/6
	definition	the sample		exposure		/1	for entire	/1		
	/1	/1					sample			
							/1			
Gander, 2008	1	1	2	0	0	0	1	1	2	4

<sup>&</sup>lt;sup>a</sup>Assessed using the Newcastle-Ottawa Quality Assessment Scale, adapted by the authors to be suitable to the non-comparative design

<sup>&</sup>lt;sup>b</sup>An overall score of 5 to 6 stars is considered as low risk of bias, 3 to 4 as unclear risk of bias, and 2 or less as high risk of bias 

Supplementary table 8. Physician health and wellness outcomes and associations with fatigue

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome	
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and		
		time points		time points		
Surgeons						
Jackson, 2017	CS	Not feeling well rested: self-	71% healthy, 28% unhealthy in	Job satisfaction: Abridged	Job satisfaction in those more vs. less satisfied:	
		reported as 'unhealthy'	terms of being well rested	Job in General Scale;	Healthy (well rested): 85% vs. 58%, p<0001;	
RoB: high				grouped into more or less	Unhealthy (not well rested): 15% vs. 42%, p<0.00	
		Time points NR		satisfied using the median		
				Time points NR		
Nishimura, 2014	CS	Sleep hours/night: self-	Mean±SD sleep: 5.94±1.08h	Burnout: Japanese MBI	1) Mean±SD sleep for not burned out vs. mild to	
		reported (continuous)		(severe: EE >4.0 and either	moderate vs. severe: 6.07±1.15 vs. 5.88±0.94 vs.	
RoB: unclear				DP >2.6 or PE <4.17)	5.63±0.94, p<0.05;	
		Time points NR			2) Association between sleep and burnout (OR	
				Time points NR	(95% CI)): bivariate 0.67 (0.61-0.73), p<0.001;	
					multivariate including work characteristics and	
					mental health: 0.84 (0.75-0.94), p=0.002.	
Sargent, 2009	CS	Sleep deprivation: self-	21% none, 48% a little, 23%	Burnout: MBI (norms NR);	1) Positive correlation between sleep deprivation	
		reported on a 4-point scale	quite a bit, 8% a lot	Marital satisfaction: RDAS;	and EE, DP, psychological distress, lower marital	
RoB: high		(none, a little, quite a bit, a		Psychological morbidity:	satisfaction, all p<0.001. No relationship with PA.	
		lot)		GHQ-12 score ≥4		
a		Time points NR		Time points NR		
Anesthesiologists <sup>a</sup>						
Lederer, 2006	BA	24-h shift with on-call duty;	Mean±SD sleep: 4.1±1.7h;	Stress during duty: 4-point	1) Mean stress score during duty: 2.1.	
		Sleep hours and	Number of interruptions:	scale from 'calm' to 'very		
RoB: high		interruptions: self-reported;	0.8±1.1;	demanding'		
		Tiredness: VAS from 0 (low)	Tiredness pre- vs. post-duty:			
		to 100 (high)	30.9±27.5 vs. 59.5±18.9,	Assessed post-duty		
			p=0.01.			
		Assessed pre- and post-duty				
Leitchtfried, 2011	TS	24-h shift;	ESS (mean (range)): 7.4 (4-12);	aMT6-s: urinalysis	1) aMT6-s over shift, mean (95% CI): higher at	
		Sleepiness: ESS (range: 0-	Mean±SD sleep hours:		11:00AM pre- (12.2 (6.3-8.1)) and post-shift (9.3	
RoB: high		24);	1) pre-study: 7.74±1.35h;	Assessed at 4-h intervals	(3.7-14.9)) vs. during, p=0.016;	
		Sleep hours: self-reported	2) Pre-24-h shift (11h00 on day	from 07:00 to 11:00	2) Correlations between sleep and aMT6-s (data	

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome	
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and		
		time points		time points		
		(continuous)	1: 0.13±0.35h, 19:00 on day 1:		NR): mild for sleep duration the night prior with	
			6.99±0.68h);		aMT6-s at 3PM the following day; sleep on night 2	
		Sleepiness assessed pre-shit,	3) During the 24-h shift (07h00		with aMT6-s at 3PM the next day; total sleep with	
		sleep hours pre, during and	on day 2: 0.0±0.0h, 19h00 on		aMT6-s at 11AM on third day; moderate for sleep	
		post-shift	day 2, 5.49±1.95h);		on first night with aMT6-s at 7AM and 11AM pre-	
			4) Post-24-h shift (11h00 on		shift, 11PM during 24-h shift and 11AM post-shift;	
			day 3: 0.5±0.71h, 19h00 on		total sleep pre-shift and nocturnal sleep during 24-	
			day 3: 7.06±1.18h).		h shift with aMT6-s at 11PM during shift; total	
					sleep with aMT6-s at 3PM on first and second day,	
					11PM on second day;	
					3) Correlations between ESS and aMT6-s:	
					moderate for aMT6-s at 7AM during shift, 11AM	
					on day off.	
Beaujouan, 2005	CS	Sleep deprivation: 4-point	48.8% always or frequently	Substance abuse: 93-item	1) 60.6% with drug dependence vs. 46.0% of those	
		scale (always, frequently,	feel sleep deprived	addiction and substance	without reported sleep difficulties, p<0.001.	
RoB: high		rarely, never)		abuse questionnaire	2) OR (95% CI) of addiction for frequently/always	
					vs. rarely/never sleep deprived: tobacco 1.42	
		Time points NR		Time points NR	(1.04-1.94); tranquilizer/hypnotics 3.26 (2.12-	
					5.02).	
Doppia, 2011	CS	Insufficient sleep: 4-point	28.9% reported insufficient	Burnout: CBI (mild: 1-2.4,	1) Frequency of burnout by response for sleep	
		scale (no, not really, sort of,	sleep during work time	moderate: 2.5-3.5, severe:	sufficiency: 47.6% for no/not really, 16.3% for sort	
RoB: low		yes)		3.6-5)	of/yes, p<0.001.	
		Time points NR		Time points NR		
Lindfors, 2006	CS	Sleep hours/day: self-	Sleep hours (mean (range)): 7	Stress: MOSQ on a 3-point	1) Sleep sufficiency predicted stress symptoms:	
		reported to the nearest	(5-9)	scale (no, to some extent,	bivariate $\beta$ =-0.362, p<0.001; multivariate including	
RoB: low		0.5h;		clearly);	gender, sick leave, suicide $\beta$ =-0.269, p<0.00;	
		Adequacy of sleep and rest:		Thoughts of suicide: 4-point	2) Sleep disturbance associated with thoughts of	
		self-reported (yes/no)		scale ('never' to 'have tried')	suicide, p=0.009.	
		Time points NR		Time points NR		
Saadat, 2015	CS	Sleep deprivation (<7h/24-h)	Mean±SD sleepiness on a	Simple cognitive tests: VAS	Regular day v. post-call day, mean±SD scores:	
		due to 17-h overnight shift;	regular day vs. post-call day:	from 0 (not at all) to 100	1) Simple cognitive tests: energetic 6.04±2.27 vs.	

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	
		time points		time points	
RoB: low		Sleepiness and alertness:	2.99±2.18 vs. 6.79±2.30,	(extremely);	2.53±1.87, confident 7.03±1.83 vs. 4.98±2.29,
		VAS from 0 (not at all) to	p<0.001	Mood disturbance: PMS	irritable 2.03±1.94 vs. 4.86±2.16, sleepy 2.99±2.18
		100 (extremely)		(scoring NR)	vs. 6.79±2.30, talkative 4.46±1.74 vs. 2.41±1.97, all
					p<0.001; jittery 1.44±1.74 vs. 3.12±2.34, p=0.003;
		All assessed on a regular day		All assessed on a regular day	anxiousness ns;
		and a post-call day		and a post-call day	2) PMS: tension 13.48±2.71 vs. 15.43±4.46,
					p=0.049; anger 15.24±4.41 vs. 18.14±5.92,
					p=0.005; fatigue 10.14±2.63 vs. 20.05±6.87,
					p<0.001; confusion 10.57±1.69 vs. 12.57±4.24,
					p=0.025; vigor 24.05±6.75 vs.16.67±5.70, p<0.001;
					depression: ns; total mood disturbance:
					42.57±15.26 vs. 70.90±6.91, p<0.001.
ER or ICU physician	าร				
Dutheil, 2013	RCT	14-h or 24-h shift;	1) Sleep duration and quality	Stress: VAS from 0 (low) to	1) Stress: higher following 14-h and 24-h shifts vs.
		Sleep hours: self-reported	lower during shifts (14h and	100 (high);	the control day, p<0.05 (data NR);
RoB: high		sleep and wake time;	24h) than any other day, and	IL-8: urinalysis	2) IL-8: higher following 24-h shift vs. control
		Sleep quality: VAS from 1	lower during the 24-h vs. 14-h		(p=0.007) and 14-h shift (p=0.015); ns difference
		(low) to 100 (high);	shift (p<0.05);	Assessed at 08:30 and 18:30	between 14-h shift and control day;
		Mental and physical fatigue:	2) Mental and physical fatigue	on each day of protocol	3) Correlations with IL-8: sleep hours pre-24-h
		VAS from 1 (low) to 100	higher after 14-h and 24-h shift		shift, r=-0.627, p=0.007; poor sleep quality during
		(high)	vs. control day (data NR).		14-h and 24-h shifts, r=0.452, p=0.031;
					4) Multivariable regression: 24-h shift increased IL-
		Assessed on day prior to			8 by 1.9ng vs. control day, p=0.007; ns association
		shift; during shift; each day			with 14-h shift, mental or physical fatigue, sleep
		of protocol (work, off,			deprivation, 14-h shift.
		clerical, control)			
Sende, 2012	CS	Fatigue and sleep	NR	Most important sources of	1) 78% indicated that sleep loss and fatigue were
		deprivation as sources of		stress among 4 categories	sources of stress.
RoB: high		stress		(work-related, patient-	
				related, organizational,	
		Time points NR		individual)	
				Time points NR	

Study	Study	<b>Exposures or interventions</b>		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	
		time points		time points	

Generalists <sup>b</sup>					
Harbeck, 2015 RoB: unclear	CS	24-hours on-call shift with sleep disturbance: self-reported number of sleep disturbances and hours of sleep per night  Assessed before a normal day shift, and after a 24-h on call shift	1) Sleep hours on a normal day vs. following a 24-h shift: <2 hours: 0 vs. 5.9%; 2-4 hours: 5.9% vs. 47.1%; 4-6 hours: 11.8% vs. 35.3%; >6 hours: 82.4% vs. 11.8% 2) Number of sleep disturbances a normal day vs. following a 24-h shift: 0: 82.4% vs. 11.8%; 1: 11.8% vs. 35.3%; 2: 5.9% vs. 47.1%; 3: 0% vs. 5.9%; 4: 0% vs. 0%; >4: 0% vs. 0%	Biochemical (laboratory values) and physiological (heart rate variability, skin resistance, blood pressure) stress parameters  Assessed before a normal day shift, and after a 24-h on call shift	Before a normal shift vs. after overnight call shift:  1) Biochemical parameters: no changes in any parameter except for thyroid stimulating hormone which was higher after the on-call shift (p = 0.049, data NR);  2) Physiological parameters: no significant change in any parameter
Pit, 2014  RoB: unclear	CS	Work-related sleep disturbance: 7-point scale from 'never' to 'every day' Time points NR	Work-related sleep disturbance: 41% never, 59% a few times a year to every day	Early retirement (<65 years) intentions (yes/no)  Time points NR	For sleep disturbance a few times a year to every day vs. never:  1) Intention to retire early: 74% vs. 26%, p<0.01;  2) Association with intention to retire early (OR (95% CI)): univariate 3.6 (1.47-8.80), p<0.01;  multivariate including work, occupational, individual factors 2.91 (1.11-7.6), p<0.05;  4) RR (95% CI) for intention to retire early: 2.0 (1.18-3.49); attributable fraction: 50.0%; population attributable fraction: 37.1%.
Pit, 2016 RoB: unclear	CS	Work-related sleep disturbance: 7-point scale from 'never' to 'every day' Time points NR	Work-related sleep disturbance: 41% never, 59% a few times a year to every day	Sickness presenteeism: 'yes' response indicated 1 or more days  Assessed for the past 12 months	For sleep disturbance a few times a year to every day vs. never:  1) Sickness presenteeism: 32% vs. 68%, p=0.018; 2) Association with sickness presenteeism (OR (95% CI)): 2.92 (1.19-7.16), p=0.02.

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome	
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and		
		time points		time points		
Roberts, 2014	CS	Fatigue: LAS from 0 (low) to	Mean (SD) score: 5.8 (2.4) for	Impact of fatigue on daily	1) 8.7% of hospitalists and 4.3% of outpatient	
		10 (high)	hospitalists; 5.9 (2.4) for	activities (falling asleep	general internists had fallen asleep while driving	
RoB: unclear			general internists	while driving) (yes/no)	due to fatigue.	
		Assessed for the past week				
				Time points NR		
Vela-Bueno, 2008	CS	Sleep Quality: PSQI	Prevalence (% (95% CI)):	Burnout: PBM with a 7-point	Low vs. high burnout, mean±SD:	
		(Spanish): score ≥5 indicates	1) Sleep-onset latency >30	scale from 1 (never) to 7	1) Global PSQI: 2.72±2.22 vs. 7.24±4.17, p<0.001;	
RoB: low		low quality (range; 0 to 21);	minutes: 8.4 (4.8-11.9);	(always)	2) PSQI subscores: sleep quality: 0.54±0.57 vs.	
		Insomnia: DSM-IV criteria	2) Wake time after sleep onset		1.40±0.83, p<0.001; sleep latency: 0.51±0.80 vs.	
			>30 minutes: 15.4 (10.8-19.9);	Time points NR	1.38±1.03, p=0.002; sleep duration: 0.45±0.64 vs.	
		Time points NR; insomnia	3) Early morning awakening:		1.16±0.92, p=0.003; sleep efficiency: 0.21±0.57 vs.	
		symptoms in past month	22.5 (19.5-30.4);		0.77±0.98, p=0.018; sleep disturbance: ns; use of	
			4) Nonrestorative sleep: 22.5		medication: 0.14±0.49 vs. 0.57±0.83, p=0.032;	
			(17.2-27.7);		daytime dysfunction: 0.52±0.73 vs. 1.57±0.88,	
			5) Daytime impairment for ≥5		p=0.002.	
			days in past month: 14.2 (9.7-		3) Prevalence (95% CI) of insomnia symptoms:	
			18.6);		sleep latency: 5.5% (2.5-11.5%) vs. 21.1% (10.5-	
			6) Insomnia: 18.8 (13.8-23.7).		31.6%), p=0.015; wake time >30 min after sleep	
					onset: 9.4% (1.6-17.1%) vs. 25.5% (14.2-37.7%),	
					p=0.029; early awakening: 14.5% (5.1-23.8%) vs.	
					45.6 (32.7-58.4%), p<0.001; somewhat/very	
					dissatisfied with sleep: 5.5% (2.5-11.5%) vs. 50%	
					(37.1-62.8%), p<0.001; day impairment: 5.5% (2.5-	
					11.5%) vs. 38.2% (25.6-50.7%), p<0.001; insomnia:	
					7.3% (0.4-14%) vs. 39.7% (27.1-52.2%), p<0.001.	
Oncologists						
Shanafelt, 2005	CS	Fatigue: LASA QOL ≤7;	75% had a high level of fatigue;	Wellbeing: 10-item LASA	1) Sleep deprivation for high vs. low overall well-	
		Sleep deprivation: 10-point	Mean±SD sleep score:	QOL, high ≥8 vs. low ≤7	being (mean±SD): 3.9±2.57 vs. 5.1±2.60, p=0.0004	
RoB: unclear		Likert scale from 0 (not at	4.5±2.65.		2) Lower fatigue predicted overall wellbeing in a	
		all) to 10 (stressful as can		Time points NR	multivariate model including personal and	
		be)			professional characteristics, p=0.002.	
		Time points NR				

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and time points	Baseline	Assessment measure and time points	
Shanafelt, 2014	CS	Fatigue: 10-point LAS (lower	Mean±SD fatigue score:	Satisfaction with WLB: 5-	1) OR (95%CI) of lower satisfaction predicted by
		scores indicate greater	5.7±2.4	point Likert scale from	high fatigue (vs. not) in multivariate model
RoB: unclear		fatigue)		'strongly agree' to 'strongly disagree'	including personal and work-related factors, and burnout: 0.489 (0.337-0.710), p<0.001.
		Time points NR		a.546. CC	2424. c. 163 (c. 163 / c. 1, 20), p
				Time points NR	
Mixed groups of p	hysicians				
Aziz, 2004	CS	Working while fatigued: 5-	NR	Stress: 47-item	1) Sources of stress: working while fatigued had a
		point scale from 'extreme'		questionnaire with a 5-point	mean±SD score of 2.44±1.20, factor loading:
RoB: high		to 'a little'		scale from 'extreme' to 'a	0.653, in factor analysis;
				little'	2) Inverse correlation between stress and working
		Time points NR			while fatigued: r=-0.270 (significance level NR).
				Time points NR	
Chen, 2008	CS	Sleepiness: ESS score ≥11	Mean±SD ESS score: 7.8±4.0,	Impact on work and	1) Impact score correlated with ESS, r=0.31,
			range: 0-20, 23% had scores	personal life: Impact	p<0.05;
RoB: high		Time points NR	≥11.	Questionnaire with a 5-	2) ESS score was higher among physicians who
				point Likert scale from 1	agree/strongly agree vs. other response: worried
				(strongly agree) to 5	about having a car accident while driving home
				(strongly disagree)	post-call: 5.4 vs. 7.0, p<0.001; sleep loss has a
					major impact on personal life: 8.4 vs. 7.0, p=0.01;
				Time points NR	3) Higher ESS scores predicted by impact score in
					multivariate regression including personal and
					work-related factors: β=0.11, p=0.005.
Elovaino, 2015	CS	Sleeping problems: Jenkins	Mean±SD score:	Jobs demands: 5 items	There was no association between sleeping
		Scale with a 6-point scale	2006: 2.30 (1.00);	scored on a 5-point scale	problems in 2006 and job demands or control in
RoB: low		from 1 (never) to 6 (every	2010: 2.35 (1.05).	from 1 (strongly disagree) to	2010.
		night)		5 (strongly agree);	
				Job control: 3 items derived	
		Assessed in 2006 and 2010		from the Karasek Job	
				Questionnaire	
Heponiemi, 2014	CS	Sleeping problems: Jenkins	Mean±SD (range) score:	Psychological distress: GHQ-	1) Sleeping problems associated with job
		Scale <sup>81</sup> with a 6-point scale	2.30±1.00 (1-6)	12 with a 4-point scale (low	satisfaction, $\beta$ =-0.12, p<0.001, psychological
RoB: low		from 1 (never) to 6 (every		to high);	distress, β=0.18, p<0.001;

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome	
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and		
		time points		time points		
		night)		Job satisfaction: JDS with a	2) Total indirect effect of on-call duty through two	
				Likert scale from 1 (strongly	mediators (sleeping problems, work interference	
		Assessed in 2006		disagree) to 5 (strongly	with family) (R <sup>2</sup> (95% CI)): job satisfaction 0.06 (-	
				agree)	0.059, -0.016), p<0.001; psychological distress 0.10 (0.023, 0.081), p<0.001.	
				Assessed in 2010	(0.023, 0.002), p (0.001.	
Mahmood, 2016	CS	Sleep deprivation: self-	Mean±SD hours:	Alcohol use disorders:	There was no association between hours of sleep	
		reported mean hours of	4 years: 4.52 (2.79);	Modified 9-item version of	when on call and hazardous drinking behaviours	
RoB: high		sleep when on call	10 years: 5.38 (6.36);	the Alcohol Use Disorder	(p=0.732)	
-		·	15 years: 6.41 (7.14).	Identification Test (AUDIT)		
		Assessed at 4 years, 10		≥6 for men and ≥5 for		
		years, and 15 years post-		women.		
		graduation				
				Assessed at 4 years, 10		
				years, and 15 years post-		
				graduation		
Shirom, 2010	CS	Tiredness and exhaustion:	NR	Burnout: SMBM on a 7-	1) Correlation between physical fatigue subscale	
		SMBM Physician Fatigue		point scale from 1 (almost	and overall burnout: 0.88, p<0.05;	
RoB: low		Subscale on a 7-point scale		never) to 7 (always)	2) In a predictive structural model for burnout,	
		from 1 (almost never) to 7			physical fatigue accounted for unique variance in	
		(always)			the burnout items, not accounted for by total	
					burnout (R <sup>2</sup> =0.24).	
		Time points NR				
Smith, 2017	CS	Sleep deprivation: self-	NR	Mental and physical illness:	Some physicians reported developing mental	
		reported via open-ended		self-reported via open-	illness (e.g., bipolar disorder, alcohol misuse) due	
RoB: unclear		comments		ended comments	to tiredness and stress at work; others developed	
					physical health problems due to sleep deprivation	
		Time points NR		Time points NR	poor eating habits and lack of exercise.	
Starmer, 2016	CS	Sleep deprivation: <7 hours	27.7% sleep deprived	Burnout, satisfaction with	≥7-h vs. <7-h sleep:	
		sleep in a typical 24-h period		career and life, balanced	1) Burnout (% strongly agree/agree): 26.4% vs.	
RoB: low		(self-reported)		personal and professional	39.6%, p<0.05; career satisfaction (% strongly	
				commitments: Each on a 5-	agree/agree): ns; life satisfaction (%	
		Time points NR		point Likert scale (strongly	completely/very satisfied): 76.4% vs. 55.9%,	

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome	
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and		
		time points		time points		
				agree to strongly disagree)	p<0.05; balanced personal and professional	
					commitments (% completely/very satisfied): 49.7%	
				Time points NR	vs. 26.1%.	
					2) <7-h sleep (vs. ≥7-h) (OR, 95% CI) associated	
					with life satisfaction 0.44 (0.29-0.67), p<0.05;	
					balanced personal/professional commitments 0.46	
					(0.31-0.71), p $\leq$ 0.05, in a model including work and	
			>		personal factors.	
Tokuda, 2009	CS	Sleep hours/day: self-	Mean±SD (range) sleep	Burnout: MBI (Japanese)	Maximum likelihood estimates±SE:	
		reported (continuous)	hours/day: 6±0.9 (3-8)	with a 7-point Likert	1) Sleeping time to job satisfaction: group	
RoB: low				scale: 0 (none) to 6 (every	0.990±0.458, p=0.031; ns for men; women	
		Time points NR (included		day);	1.711±0.805, p=0.034;	
		weekday and weekends)		Job satisfaction: JHPSS <sup>86</sup>	2) Sleeping time to EE: group -0.219 ±0.070,	
				with a 5-point Likert	p=0.002; men -0.215±0.082, p=0.009; ns for	
				scale: 1 (strongly	women.	
				disagree) to 5 (strongly		
				agree)		
				Time points NR		
Wada, 2010	CS	Sleep hours/day: Self-	<5 hours: 8.7% men, 9.9%	Depression: QIDS-SR;	1) Sleep hours for those with vs. without	
		reported (continuous)	women; 5 to <6 hours: 32.3%	Japanese score <5 (no	depressive symptoms: <5: 18.7% vs. 7.7% men,	
RoB: unclear			men, 34.6% women; 6 to <7	symptoms) to >20 (very	20.5% vs. 8.7% women; 5 to <6: 33.7% vs. 32.2%	
		Assessed for past month	hours: 46.0% men, 43.7%	severe symptoms)	men, 38.6% vs. 34.2% women; 6 to <7: 35.1% vs.	
		when not completing	women; ≥7 hours: 13.0% men,		46.9% men; 31.8% vs. 45.1% women;	
		overnight work	11.8% women.	Assessed for past 7 days	2) Association between <5h sleep (vs. 6-7h) and	
					depressive symptoms (OR (95% CI)): univariate	
					2.79 (1.96-3.95) for men, 2.65 (1.47-4.78) for	
					women; multivariate (including age and workload	
					factors) 2.70 (1.82-4.03) for men, 2.38 (1.11-5.10)	
					for women.	

<sup>&</sup>lt;sup>a</sup>Includes studies of anesthetists, where these were physicians.

<sup>&</sup>lt;sup>b</sup>Includes primary care physicians, internal medicine physicians, and general practitioners.

AM: morning; aMT6-s: melatonin metabolite; BA: before-after; CI: confidence interval; CBI: Copenhagen Burnout Inventory; CS: cross-sectional; DP: depersonalization; DSM: Diagnostic and Statistical Manual of Mental Disorders; EE: emotional exhaustion; ER: emergency; ESS: Epworth Sleepiness Scale; GHQ: General Health Questionnaire; h: hour(s); #HPSS: ..

.ed Occupationa.

.c; PBM: Pines Burnout i,

pressive Scale – Self-Reported; i,

un; SE: standard error; SMBM: Shirom-Iv.

e ICU: intensive care unit; IL-8: interleukin-8; JDS: Job Diagnostic Survey; JHPSS: Japanese Hospital Physicians Satisfaction Scale; LAS: linear analog scale; LASA: linear analog assessment scales; MBI: Maslach Burnout Inventory; MOSQ: Modified Occupational Stress Questionnaire; min: minute(s); NA: not applicable; NR: not reported; ns: not statistically significant; OR: odds ratio; PA: personal achievement; PBM: Pines Burnout Measure; PE: professional efficacy; PM: afternoon; PMS: Profile of Mood States; PSQI: Pittsburgh Sleep Quality Index; QIDS-SR: Quick Inventory Depressive Scale - Self-Reported; QOL: Quality of Life; RCT: randomized controlled trial; RDAS: Revised Dyadic Adjustment Scale; RoB: Risk of Bias; SD: standard deviation; SE: standard error; SMBM: Shirom-Melamed Burnout Measure; TS: time series; US: United States of America; VAS: visual analog scale; vs.: versus; WLB: work-life balance

Supplementary table 9. Performance and safety outcomes related to fatigue or sleep loss among physicians in independent practice

Study	Study	Exposures or intervention		Outcomes	Associations between exposure and outcome	
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and		
		time points	time points			
Surgeons						
Uchal, 2005	RCT	Sleep deprivation from a 24-	Median (range) sleep hours:	Surgical performance:	Post call vs. post-work:	
		h call shift vs. 8-h work;	1.5 (0-3) post-call vs. 6.5 (5-	laparoscopic surgical	1) Product quality: no difference in accuracy	
RoB: unclear		Sleep hours: self-reported	9) post-work, p<0.05;	simulator for product quality,	error, tissue damage, leak rate;	
		(continuous);	Median ESS score: 7.0 post-	procedure effectiveness	2) Procedure effectiveness: no difference in	
		Sleepiness: ESS (moderate:	call vs. 5.5 post-work, ns.		goal-directed actions, non-goal directed	
		10-15, severe: ≥16)		Assessed post-call and post-	actions, operating time.	
				work		
		Assessed post-call and post-				
		work				
Chu, 2011	СО	Sleep deprivation: self-	Of 4,047 procedures, 83	Surgical performance: CABG,	For 0-3 vs. 3-6 vs. >6 hours of sleep: no	
		reported hours, moderate	(2.1%) performed by	ACC	difference in CABG or ACC.	
RoB: low		(3-6h) or severe (<3h)	severely sleep-deprived and			
			1,595 (39.4%) moderately	Assessed during surgery		
		Assessed the night before	sleep-deprived surgeons			
		surgery				
Ellman, 2004	СО	Sleep deprivation:	Of 6,751 procedures, 339	Surgical performance: CABG,	Sleep deprived vs. non-sleep deprived: no	
		performed a case starting	(5%) performed by sleep-	ACC	difference in CABG or ACC.	
RoB: low		22:00 to 05:00, or ending	deprived surgeons			
		22:00 to 07:30 and another		Assessed during surgery		
		case in the next 24-h		96		
Govindarajan,	CO	Sleep deprivation: treated	NR	Surgical performance:	Sleep deprived vs. non-sleep deprived: no	
2015		patients from midnight to		duration of surgery	difference in duration of surgery, even after	
		07:00 and performed a			stratification by type of procedure.	
RoB: low		subsequent case on the				
		same day				
Amirian, 2014	BA	17-h night call shift;	Naps pre-call: 11 (37%)	Surgical performance:	Pre- vs. post-call:	
		Sleep hours during the shift:	napped for median (IQR) 90	LapSimGyn laparoscopic	1) LapSimGyn: no difference in total time, blood	
RoB: high		Wrist-mounted Micro-Mini-	(58-128) min;	simulation for time, blood	loss, instrument path length, instrument	
		Motionlogger;	Median (IQR) sleep: 91 (62-	loss, instrument path;	angular path; napping did not affect	
		Sleepiness: KSS	123) min on the pre-call	D2 test of attention and	performance;	
			night vs. 430 (329-449) on	concentration	2) D2 test: improvement in concentration,	

Study Risk of Bias (RoB)	Study design	Exposures or intervention		Outcomes	Associations between exposure and outcome
		Assessment measure and time points	Baseline	Assessment measure and time points	
		Assessed on pre-call and on-	the on-call night, p<0.001;		p<0.05. No changes in any other parameters;
		call day; sleepiness assessed	Sleep on-call: 12 (40%) slept	Assessed on pre-call and on-	3) ns difference in laparoscopic simulation time
		during shift	for median (IQR) 98 (39-135)	call day	in those who slept during the shift vs. not.
			min;		
			Significant development of		
			sleepiness during shift		
			(p<0.001), plateau score of 7 at 04:00 to 08:00.		
Gerdes, 2008	BA	On-call shift;	Fatigue differential from pre-	Psychomotor performance:	1) Pre- to post-call: decrease in all measures of
deraes, 2000	D/ (	Fatigue: questionnaire	to post-call (range): 1-7	virtual ring transfer task for	psychomotor proficiency (p<0.05, data NR)
RoB: high		designed by Behrenz &	(units unclear);	gesture-level proficiency,	except elapsed time; no change in number of
- 0		Monga, 1999;	Sleep during call (range): 1-	hand movement smoothness,	psychomotor errors; increase cognitive errors
		Sleep hours: self-reported	5h	tool movement smoothness,	(p<0.05, data NR);
		(continuous)		elapsed time	2) Cognitive errors increased exponentially as
					fatigue ratings increased (R <sup>2</sup> =0.9219) and as
		Assessed in 3 sessions pre-		Assessed in 3 sessions pre-	hours of sleep declined (R <sup>2</sup> =0.933).
		and post-call		and post-call	
Shanafelt, 2010	CS	Degree of fatigue as a	NR	Perceived recent major	1) Prevalence of perceived recent major
		contributor to errors (self-		medical errors (self-reported)	medical error: 8.9%;
RoB: high		reported)			2) Of those reporting an error, 6.9% listed
				Assessed for the past 3	degree of fatigue as the greatest contributing
		Assessed for the past 3		months	factor.
A		months			
Anesthesiologists					
Lederer, 2006	BA	24-h shift, on-call duty;	Mean±SD sleep: 4.1±1.7h;	Psychomotor performance:	Pre- vs. post-duty, mean±SD:
D D I ' I		Sleep hours and	Number of interruptions:	reaction time, critical flicker	1) Psychometric testing: recognition reaction
RoB: high		interruptions: self-reported;	0.8±1.1;	fusion, response measure,	time (ms): 439.6±50.8 vs. 480.3±58.9; motor
		Tiredness: VAS from 0 (low)	Tiredness pre- vs. post-duty:	peripheral awareness;	reaction time (ms): 252.8±39.3 vs. 465.4±65.0;
		to 100 (high)	30.9±27.5 vs. 59.5±18.9, p=0.01.	Concentration ability: scale of 0 (low tiredness) to 100	total reaction time (ms): 690.8±73.4 vs. 746.5±113.7; critical flicker fusion (Hz): 29.0±2.3
		Assessed pre- and post-duty	μ-0.01.	(maximum tiredness)	vs. 28.7±3.7; cruical nicker fusion (Hz): 29.0±2.3
		7.55c55ca pre and post-duty		(maximum diculiess)	647.8±126.7 vs. 598.3±138.1, peripheral
				Assessed pre- and post-duty	awareness task recognition time: 58.9±59.2 vs.
				. issessed pre-una post duty	2.1.2. 2.1.2.2 task recognition time: 30.3±33.2 vs.

Study Risk of Bias (RoB)	Study design	Exposures or intervention		Outcomes	Associations between exposure and outcome
		Assessment measure and	Baseline	Assessment measure and time points	
		time points			
					51.6±47.5;
					2) Concentration ability: 26.4±23.5 vs.
					56.3±23.0, p=0.007.
Chang, 2013	CS	15-h in-house overnight call;	Median (IQR) ESS: 9 (9), 64%	Psychomotor performance:	1) Afternoon baseline vs. pre-call: no difference
		Sleepiness pre-call: ESS ≥9;	scored ≥9;	reaction time; CCPT II; N-	in reaction time, CCPT, N-back, of HVLT;
RoB: unclear		Sleep hours: self-reported	Median (IQR) hours slept	back; HVLT (3 trials of 12	Morning baseline vs. post-call:
		(continuous)	during shift: 1 (0-3).	words)	1) No change in auditory or visual reaction time
					2) CCPT (t-scores): No change in detectability,
		Sleepiness assessed pre-call,		Assessed at baseline and pre-	response style, hit reaction time,
		sleep hours during call		and post-call	omissions/commissions;
					3) N-back % accuracy: no change for auditory,
					visual, or mean N-value;
					4) HVLT (t-score): mean for trials 1-3: 48.6±7.6
					vs. 41.5±9.9 (p=0.04); delayed recall: ns;
					5) No correlation between ESS scores pre-call o
					sleep during shift and any measure of
					psychomotor performance.
Gander, 2000	CS	Nights of work-related sleep	NR	Risk of fatigue-related errors:	1) Risk of fatigue-related errors increased with
		disturbance: self-reported		questionnaire modelled after	increasing nights of work-related sleep
RoB: low		(continuous)		Gravenstein et al., 1990	disturbance: RR: 1.25, 95% CI: 1.06-1.49.
		Assessed for the past 6		Assessed for the past 6	
		months		months	
Saadat, 2017	CS	Sleep deprivation due to an	NR	Reaction time: PVT	Mean (SD) reaction time was slower post-call
		overnight call shift			(297.76 (83.75)) vs. on a regular day (266.58
RoB: low				Assessed after an overnight	(38.35)), p=0.047.
				call shift and the morning of a	
				regular (non-call) day	
Gander, 2008	NC	Sleep loss across consecutive	≥2 hours sleep <baseline: 8%<="" td=""><td>Psychomotor performance:</td><td>1) In fixed model analysis for reaction time</td></baseline:>	Psychomotor performance:	1) In fixed model analysis for reaction time
		working days or on-call	of 24-h periods that included	PVT	including sleep, time since waking, work hours:
RoB: unclear		work: Wrist-mounted	day work vs. 14% that		acute sleep loss associated with slower median
		Actiwatch (Mini Mitter,	included day + call;	Assessed within 2 hours pre-	reaction time, F <sub>(1.184)</sub> =5.70, p<0.05; longer time
		Bend, Oregon, US), sleep	Sleep hours: mean 0.6h less	and post-call	since waking associated with poorer

Study Risk of Bias (RoB)	Study design	Exposures or intervention		Outcomes	Associations between exposure and outcome
		Assessment measure and	Baseline	Assessment measure and	
		time points		time points	
		and duty diary	sleep when working day		performance on the slowest 10%, $F_{(1,185)}$ =5.13,
			shifts (p=0.014) and 0.8h less		p<0.05;
		Assessed over a 2-week	sleep when working day		2) Reaction time across 12 consecutive work
		period including a weekend	shifts + call (p=0.013) vs. off.		days: no change in pre-duty reaction times but
		of rostered shifts or on-call			post-duty reaction times slowed linearly,
					median -3.38, p<0.001; decline in performance
					across 10 minutes became progressively
					steeper both pre- and post-duty, p=0.020.
ER or ICU physicians	S				
Sanches, 2015	CS	Acute sleep deprivation (<5h	Non-sleep deprived vs. sleep	Psychomotor performance	Sleep deprived group vs. non-sleep deprived,
		of night sleep after a night	deprived:	via Battery Test Reaction 5	mean±SD:
RoB: unclear		shift of 12h)		(v1): StimulTest, InstrucTest,	1) InstrucTest: correct answers: 169.4 (16.0) vs
		Sleep hours: 7-day	PSQI >5: 0% vs. 33%, ns;	MovemTest; TP test of visual	148.3 (28.3), p=0.070; wrong answers: ns;
		Actigraphy via SenseWear®	ESS≥10: 11% vs. 67%	attention	perfection index (%): 99.6 (0.3) vs. 98.9 (1.3),
		Pro2 Armband;	Sleep time (mean±SD) in		p=0.021; response latency (sec/click): ns;
		Sleepiness: ESS;	week before tests: duration	Assessed on morning after	2) StimulTest: correct answers: 170.7 (21.9) vs.
		Sleep quality: PSQI	and number of naps higher	night shift 8	145.1 (17.9), p=0.022; wrong answers: ns;
			in sleep deprived group, but		perfection index (%): ns; response latency
		Assessed the week and night	diurnal sleep hours lower,		(sec/click): 1.06 (0.1) vs. 1.24 (0.1), p=0.022;
		before the psychomotor	428.6±30.1 vs. 375.8±55.9,		3) MovemTest: ns for any parameter;
		tests	p=0.038;		4) TP: omitted symbols: 34.2±18.4 vs.
			Sleep quality (mean±SD):		62.7±44.0, p=0.034; concentration index (%):
			week before tests: 3.3±0.7		14.1±8.9 vs. 30.0±25.9, p=0.019; quality index
			vs. 2.6±0.3, p=0.013;		(%): 13.8±8.6 vs. 29.2±26.4, p=0.031;
			night before tests: 3.1±0.8		correct/wrong symbols: ns;
			vs. 1.9±1.0, p=0.020.		Correlations between sleep and tests:
					1) TP for sleep hours nights 1-6: omitted
					symbols: r=-0.686, p=0.011 for non-sleep-
					deprived, ns for sleep-deprived; concentration
					index (%): r=-0.359, p=0.037 for sleep-deprived
					ns for non-sleep deprived; r=-0.359, p=0.037 fo
					the group; no other significant correlations;
					2) No correlation between PSQI, ESS and any of

Study Risk of Bias (RoB)	Study design	Exposures or intervention		Outcomes	Associations between exposure and outcome
		Assessment measure and	Baseline	Assessment measure and	
		time points		time points	
					the psychomotor tests.
Generalists <sup>b</sup>					
Harbeck, 2015	CS	24-hours on-call shift with sleep disturbance: self-reported number of sleep disturbances and hours of sleep per night  Assessed before a normal day shift, and after a 24-h on call shift	1) Sleep hours on a normal day vs. following a 24-h shift: <2 hours: 0 vs. 5.9%; 2-4 hours: 5.9% vs. 47.1%; 4-6 hours: 11.8% vs. 35.3%; >6 hours: 82.4% vs. 11.8% 2) Number of sleep disturbances a normal day vs. following a 24-h shift: 0: 82.4% vs. 11.8%; 1: 11.8% vs. 35.3%; 2: 5.9% vs. 47.1%; 3: 0% vs. 5.9%; 4: 0% vs. 0%; >4: 0% vs. 0%	Neurocognitive parameters: computerized attentional test (vigilance, alertness); D2 letter cancellation test (divided attention); Trail Making Test (visual attention, task switching); Digit Span, Digit Symbol Substitution Test, Weschler Memory Scale (memory functions)  Assessed before a normal day shift, and after a 24-h on call shift	Intrinsic alertness, focused attention and vigilance were similar on both occasions; Phasi alertness improved following the on-call shift: mean (SD) 24.8 (15.6) vs. 38.3 (21.5), p = 0.022
Mixed specialties o	r undefine	ed populations			
Chen, 2008	CS	Sleepiness: ESS score ≥11	Mean±SD ESS score: 7.8±4.0, range: 0-20, 23%	Impact on work and personal life: Impact Questionnaire	1) Impact score correlated with ESS, r=0.31, p<0.05;
RoB: high		Time points NR	had scores ≥11.	with a 5-point Likert scale from 1 (strongly agree) to 5 (strongly disagree) Time points NR	2) ESS score was higher among physicians who agree/strongly agree vs. other response: written an incorrect order: 8.8 vs. 7.3, p=0.02; might fall asleep while examining a patient: 13.2 vs. 7.7, p=0.001; look forward to sleeping at grand rounds: 10.4 vs. 7.4, p=0.002; 3) No difference in ESS score for those who agree/strongly agree vs. other response: work unaffected by sleep loss and fatigue, thinking is unaffected by sleep loss, sleep loss and fatigue affect my medical decisions, have heard of
					unaffected by sleep loss and fati unaffected by sleep loss, sleep lo

Study Risk of Bias (RoB)	Study design	Exposures or intervention		Outcomes	Associations between exposure and outcome
		Assessment measure and	Baseline	Assessment measure and	
		time points		time points	
					on post-call days, have made medical errors
					because of sleep loss and fatigue;
					4) Higher ESS scores predicted by impact score
					in multivariate regression including personal
					and work-related factors: $\beta$ =0.11, p=0.005.
Heponiemi, 2014	CS	Sleeping problems: 4-item	Mean±SD (range) score:	Work ability: Work Ability	1) On-call duty had an indirect effect on work
		Jenkins Scale on 6-point	2.30±1.00 (1-6)	Index on scale from 1 (could	ability (R <sup>2</sup> =0.11, 95% CI: -0.122, -0.031,
RoB: low		scale from 1 (never) to 6		not work at all) to 10 (best	p<0.001) through two mediators (work
		(every night)		work ability)	interference with family, sleeping problems);
					2) Sleeping problems inversely associated with
		Assessed in 2006		Assessed in 2010	work ability, $\beta$ =-0.29, p<0.001.
Kanieta, 2011	CS	Sleep hours: self-reported	Insufficient rest: 32.5%;	Self-reported medical	1) Prevalence of medical incidents (% (95% CI)):
		(continuous)	Daytime sleepiness: 3.5%;	incidents: 4-point scale from	sleep deprived (26.8% (24.2, 29.4)) vs. not
RoB: unclear		Sleepiness and sleep	Insomnia: 20.0%;	1 (never) to 4 (often)	(15.2% (13.7, 16.7)), p<0.01; insomnia (24.8%
		difficulties: 5-point scale	Sleep time (mean±SD min):		(21.6, 28.0)) insomnia vs. not (17.6% (16.2,
		from 1 (never) to 5 (always);	279.8±60.9	Assessed for the past month	19.0)), p<0.01; ≥6h sleep (18.3% (16.8, 19.8)) vs.
		Insomnia: ≥3 sleep			<6h (21.7% (18.8, 24.6)), p=0.03;
		difficulties			2) Predictors of medical incidents in
					multivariate model including personal and
		Assessed for the past month			work-related factors (OR (95% CI)): lacking rest
					due to sleep deprivation vs. not (1.65 (1.33-
					2.04)), p<0.01); insomnia vs. not (1.45 (1.16-
					1.82), p<0.01); ns for sleep hours.
Sexton, 2001	CS	Fatigue as a factor impacting	NR	Performance effectiveness	1) "When fatigued, I perform effectively during
		performance		measured by 1 question:	critical phases of operations/patient care":
RoB: high				agree, neutral, disagree	Anesthetic: 47% agree; 15% neutral; 38%
		Time points NR			disagree;
				Time points NR	Surgical: 70% agree; 12% neutral; 18% disagree.
Shirom, 2006	CS	Tiredness and exhaustion:	NR	Quality of care: Adapted 15-	1) Quality of care positively predicted by fatigue
		SMBM Physician Fatigue		item SERVQUAL with a 5-	in a model incorporating several other
RoB: low		Subscale on a 7-point scale		point Likert scale from 1 (very	components of burnout, $\beta$ =0.17, p<0.05.
		from 1 (almost never) to 7		small extent) to 5 (very large	
		(always)		extent)	

Study	Study	Exposures or intervention		Outcomes	Associations between exposure and outcome		
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and			
		time points		time points			
		Time points NR		Time points NR			
Smith, 2017	CS	Sleep deprivation: self-	NR	Perceived competence: self-	Some physicians indicated that continual		
		reported via open-ended		reported via open-ended	tiredness and exhaustion led to concerns that it		
RoB: moderate		comments		comments	would affect their competence; some felt that		
					professional performance was compromised at		
		Time points NR		Time points NR	times of physical and mental fatigue.		
Tanti, 2017	CS	Fatique: questionnaire on	NR	Prescribing errors:	Perception of the contribution of fatigue to		
		contributors to prescribing		questionnaire on contributors	prescribing errors differed by physician type		
RoB: high		errors, with a 5-point Likert		to prescribing errors, with a	(p<0.05): 34% of community doctors, 96%		
		scale (very high to very low		5-point Likert scale (very high	hospital doctors, 8% of office-working doctors		
		association)		to very low association)	perceived a very high or high association		
					between fatigue and prescribing errors.		
		Time points NR		Time points NR			

<sup>&</sup>lt;sup>a</sup>Includes studies of anesthetists, where these were physicians.

ACC: aortic cross-clamp time; BA: before-after; CABG: cardiopulmonary bypass time; CCPT II: Connor's Continuous Performance Test II; CI: confidence interval; CO: cohort; CS: cross-sectional; ER: emergency; ESS: Epworth Sleepiness Scale; h: hour(s); HVLT: Hopkin's Verbal Learning Task; Hz: Hertz; ICU: intensive care unit; IQR: interquartile range; KSS: Karolinska Sleep Scale; min: minutes; ms: millisecond(s); N-back: Dual N-back test; NA: not applicable; NR: not reported; ns: not statistically significant; OR: odds ratio; PSQI: Pittsburgh Sleep Quality Index; PVT: Psychomotor vigilance Performance Task; RR: risk ratio; RCT: randomized controlled trial; RoB: Risk of Bias; SD: standard deviation; SE: standard error; SERVQUAL: Service Quality Measure; SMBM: Shirom-Melamed Burnout Measure; TP: Toulouse-Piéron test; TS: time series; US: United States of America; vs.: versus

<sup>&</sup>lt;sup>b</sup>Includes primary care physicians, internal medicine physicians, and general practitioners.

Supplementary table 10. Patient outcomes related to fatigue or sleep restriction among physicians in independent practice

Study	Study	Exposures		Outcome Measures	Associations between exposure and outcome		
Risk of Bias (RoB)	design	Intervention or assessment	Baseline	Assessment scale and time			
		scale and time points		points			
Surgeons							
Chu, 20	СО	Sleep deprivation: moderate	Of 4,047 procedures, 83	Chart review: mortality, surgical	1) 0-3 vs. 3-6 vs. >6 hours of sleep: No		
		(3-6 h) or severe (<3-h) sleep	(2.1%) performed by	complications, length of stay	difference in incidence of mortality, incidence		
RoB: low		deprivation the night before	severely sleep-deprived,		of 10 major complications (except septicemia,		
		surgery (self-reported hours)	1,595 (39.4%) by	Assessed during and post-	3.6% vs. 0.9% vs. 0.8%, p=0.03), ICU length of		
			moderately sleep-	surgery	stay; in-hospital length of stay (days): 7.0 vs.		
			deprived surgeons		6.0 vs. 7.0, p<0.001.		
Ellman, 2004	СО	Sleep deprivation: performed	Of 6,751 procedures, 339	Chart review: mortality, surgical	1) Sleep deprived vs. non-sleep deprived: no		
		a case starting 22:00 to 05:00,	(5%) were performed by	complications, length of stay	difference in mortality, need for blood		
RoB: low		or ending 22:00 to 07:30 and	sleep deprived surgeons		products, complications (operative, neurologic,		
		performed a subsequent case		Assessed during and post-	renal, infectious, pulmonary), in-hospital		
		in the next 24-h		surgery	length of stay.		
Govindarajan,	СО	Sleep deprivation: treated	NR	Chart review: mortality, surgical	1) Sleep deprived vs. non-sleep deprived: no		
2015		patients from midnight to		complications, readmission,	difference in mortality, surgical complications,		
		07:00 and performed a		length of stay	readmissions within 30 days, or length of stay.		
RoB: low		subsequent case on the same					
		day		Assessed during and post-			
				surgery			
Rothschild, 2009	CO	Sleep deprivation: daytime	NR	Chart review: frequency of	1) Post-nighttime vs. control: no difference in		
		procedures following an		adverse surgical complications	number of procedures with complications,		
RoB: low		overnight procedure;			total number of complications, preventable		
		Sleep opportunity: 0-6h, <6h		Assessed during and post-	complications, type of complications;		
				surgery	2) Operating room procedures with		
					complications, OR (95% CI): 8.5% for 0-6h sleep		
					vs. 3.1% for >6h sleep, 2.70 (1.13-6.48),		
					p=0.03;		
					3) All procedures with complications, OR (95%		
					CI): 6.2% for 0-6h sleep vs. 3.4% for >6h sleep,		
					1.72 (1.02-2.89), p=0.04.		
Schieman, 2007	СО	Fatigue: surgeon billed for	Of 270 procedures, 22	Chart review: surgical	1) Fatigued vs. non-fatigued surgeons: no		
		clinical work after 22:00 the	(8%) were performed by	complications, length of stay,	difference in intra- or post-operative		
RoB: low		night before surgery	fatigued surgeons	mortality, cancer recurrence	complication rate, length of stay, in-hospital		

Study	Study	Exposures		Outcome Measures	Associations between exposure and outcome		
Risk of Bias (RoB)	design	Intervention or assessment	Baseline	Assessment scale and time			
		scale and time points		points			
					length of stay, cancer recurrence.		
				Assessed during and post-			
				surgery			
Vinden, 2014	СО	Sleep deprivation (at risk):	Of 94,183 surgeries,	Chart review: conversion to	1) At risk vs. not at risk surgeon: no difference		
		surgeon worked 00:00 to	2,078 (2.2%) were	open procedure (from	in incidence of conversion to open procedure,		
RoB: low		07:00 and performed surgery	performed by surgeons	laparoscopic), iatrogenic injuries,	iatrogenic injuries, mortality, in either		
		07:00 to 18:00	who were 'at risk'	mortality	univariate or multivariate analyses.		
				Assessed during and post-			
				surgery			
Obstetricians							
Rothschild, 2009	СО	Sleep deprivation: daytime	NR	Chart review: frequency of	1) Post-nighttime vs. control: no difference in		
		procedures following an		adverse obstetric complications	number of procedures with complications,		
RoB: low		overnight procedure;			total complications, preventable		
		Sleep opportunity: 0-6h, <6h		Assessed during and post-	complications, type of complications;		
				delivery	2) No association between sleep deprivation		
					and proportion of procedures with		
					complications, nor difference for 0-6h vs. >6h		
					of sleep opportunity.		

Cl: confidence interval; CO: cohort; h: hours; NR: not reported; OR: odds ratio; RoB: Risk of Bias; SD: standard deviation; US: United States of America; vs.: versus

**BMJ** Open

Page 74 of 75



Appendix 1. PRISMA 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.	4-5
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	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.	5
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.	6
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.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Supplementa file 1
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).	6
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.	6-7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7
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.	7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I <sup>2</sup> ) for each meta-analysis.	7, no meta- analysis



#### Appendix 1. PRISMA checklist

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).	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	Not applicable
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.	7, Figure 1
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.	7-10, Table 1, Supplementary file 2
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).	10, Supplementary file 3
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.	Supplementary files 4-6
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	10-15 (no meta-analysis)
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Not applicable
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Not applicable
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).	15-17
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).	17
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	17
FUNDING	ı		
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.	18

43 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(7): e1000097. 44 doi:10.1371/journal.pmed1000097

## **BMJ Open**

## The impact of fatigue and sleep restriction on physician and patient outcomes: A systematic review

Journal:	BMJ Open
Manuscript ID	bmjopen-2018-021967.R1
Article Type:	Research
Date Submitted by the Author:	27-Apr-2018
Complete List of Authors:	Gates, Michelle; University of Alberta, Pediatrics Wingert, Aireen; University of Alberta, Pediatrics Featherstone, Robin; University of Alberta, Pediatrics Samuels, Charles; Centre for Sleep and Human Performance Simon, Christopher; Canadian Medical Association Dyson, Michele; University of Alberta, Pediatrics
 <b>Primary Subject Heading</b> :	Occupational and environmental medicine
Secondary Subject Heading:	Evidence based practice, Health services research, Medical education and training, Occupational and environmental medicine
Keywords:	Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, SLEEP MEDICINE

SCHOLARONE™ Manuscripts

mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

The impact of fatigue and sleep restriction on physician and patient outcomes: A systematic review

Michelle Gates, Aireen Wingert, Robin Featherstone, Charles Samuels, Christopher Simon, Michele P Dyson

**Michelle Gates**, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-486C, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Project coordinator

**Aireen Wingert**, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-490B, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Project coordinator

**Robin Featherstone**, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-486D, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Research librarian

**Charles Samuels**, Centre for Sleep and Human Performance, Suite 106, 51 Sunpark Drive SE, Calgary, Alberta, Canada, T2X 3V4
Physician and medical director

**Christopher Simon**, Canadian Medical Association, 1867 Alta Vista Drive, Ottawa, Ontario, Canada, K1G 5W8

Advisor, physician health

Michele P Dyson, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-474, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Assistant professor

Correspondence to: MP Dyson mdyson@ualberta.ca

Word count (main text): 5,929

**Objectives:** For physicians in independent practice, we synthesized evidence on the (a) impacts of sleep restriction and fatigue on health and performance, and patient safety; (b) effectiveness of interventions targeting sleep restriction and fatigue.

**Design:** We systematically reviewed online literature. After piloting, one reviewer selected studies by title and abstract; full texts were then reviewed in duplicate. One reviewer extracted data; another verified a random 10% sample. Two reviewers assessed risk of bias. We pooled findings via meta-analysis when appropriate, or narratively.

**Data sources:** We searched Medline, Embase, PsycINFO, CINAHL and PubMed for published studies in April 2016; Medline was updated in November 2017. We searched Embase for conference proceedings, and hand-searched meeting abstracts, association and foundation websites.

**Eligibility criteria for selecting studies:** English or French language primary research studies published from 2000-2017 examining the effect of fatigue or sleep-related exposures or interventions on any outcome among physicians in independent practice and their patients.

**Results:** Of 16,154 records identified, we included 47 quantitative studies of variable quality. 28 studies showed associations between fatigue or sleep restriction and physician health and well-being outcomes. 21 studies showed no association with surgical performance, and mixed findings for psychomotor performance, work performance, and medical errors. We pooled data from six cohort studies for patient outcomes. For sleep deprived versus non-sleep deprived surgeons, we found no difference in patient mortality (n = 60,436, RR 0.98, 95% Cl 0.84 to 1.15, p = 0.82,  $I^2$  = 0%), intraoperative complications (n = 19,798, RR 1.35, 95% Cl 0.82 to 2.21), postoperative complications (n = 60,201, RR 0.99, 95% Cl 0.95 to 1.03) or length of stay (n = 50,046, MD -0.33, 95% Cl -1.03 to 0.36).

**Conclusions:** Fatigue and sleep deprivation may be associated with negative physician health outcomes. Current evidence is inadequate to inform practice recommendations.

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text and

- The review was informed by the methods outlined by Cochrane and is reported according to the
   Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.
- The review was limited by the quality of the included studies, which was often poor. We could not
  draw definitive conclusions due to methodological weaknesses and heterogeneous outcome
  measures in the included studies.
- We have focused on evidence from high income countries; our findings may not be generalizable to other settings.

The working hours of physicians have been a topic of debate for many years.[1] Beginning in the late 1980s, evidence indicating that medical resident fatigue could negatively impact their cognitive functioning and performance, resulting in an increased risk of medical error, began to accumulate.[2] In response, by the early 2000s physicians' regulatory bodies worldwide began to take action toward restricting the work hours of medical residents and ensuring adequate time for recovery between shifts.[3-5] Since their implementation in the United States by the Accreditation Council for Graduate Medical Education (ACGME), the impact of work hour regulations has been widely researched. Still, evidence for impacts on patient care, resident training and wellbeing remains equivocal.[6-9] This is likely because work hours are only one of many contributors to fatigue and physician wellbeing. In fact, the ACGME has recently reversed the 2011 changes that limited resident work hours to 16 hours per shift and the requirement for 8 hours of time off between shifts. This decision was made in favour of promoting "flexibility" for residency training program work hours and scheduling.

The focus on medical trainees has left physicians in independent practice as a relatively neglected group in research and policy. In Canada, there is no concrete regulation on the hours or patterns in which physicians choose to work.[10] In the absence of clear policies, physicians trained under traditional systems may find it difficult to work shorter hours or take more frequent breaks.[1] Indeed, more than 40% of practicing physicians in the United States work in excess of 80 hours per week.[11] While long work hours remain a cultural norm in medicine, in comparable high-risk industries (e.g., aviation), work patterns and work hours are tightly regulated.[12] The need for similar evidence-based policies in medicine has become a topic of increased interest. Exemplar of this, an evidence-based guideline for fatigue risk management in emergency medical services,[13] informed by a comprehensive set of systematic reviews, has recently been published. For physicians, it has been argued that there is a need to adapt healthcare systems and provide support in identifying the signs of fatigue and mitigating its risks.[1]

Besides potentially affecting patient outcomes, fatigue can impact the health and wellbeing of physicians themselves. Burnout, just one outcome related to fatigue, has been described as epidemic among physicians[14-16] and ultimately affects recruitment and retention of physicians both in community and acute care settings. While the effect of physician wellbeing on the sustainability of healthcare systems has recently received increased attention,[17] evidence-based solutions to burnout

Given this void, we undertook a systematic review focusing broadly on primary research relevant to the Canadian context as a fundamental starting point to examine the effects of fatigue and chronic sleep restriction on physicians in independent practice, and on interventions to combat these effects. Our review was guided by the following research questions: Among physicians in independent practice, (1) what are the impacts of fatigue and chronic sleep restriction on physician health, physician performance, and patient safety; and (2) what is the effectiveness of interventions that target fatigue and chronic sleep restriction loss, in terms of improving physician and patient outcomes?

#### **METHODS**

#### **Review conduct**

The conduct of this systematic review was guided by Cochrane standards.[22] The research team convened to plan the key research questions and methodology but did not register a formal protocol. The findings are reported in adherence with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) Statement.[23] Ethical approval was not required for this study.

#### **Patient involvement**

Patients were not involved.

#### Literature search

An information specialist developed a search strategy that included concepts related to physicians, fatigue and sleep. On 13 April 2016 we searched the following online databases with coverage in the biomedical sciences and psychology: Medline, Embase, PsycINFO, CINAHL and PubMed, limited to English and French language articles published from 2000 to 2016. We updated the Medline search in November 2017, as this database offered the highest precision. Though fatigue among physicians is not a new phenomenon,[2] we limited our search to articles published post-2000 to include studies relevant to current physician practice. Work hour limitations have existed in European countries since 1993, but

implementation in the United States (2003)[5] and Canada (2013) for residents is more recent.[24] We aimed to include studies published in this era of increased awareness about the potential impacts of long work hours. To locate unpublished studies, we searched Embase for conference proceedings since 2000 and hand-searched meeting abstracts of the Canadian Conference on Physician Health and the International Conference on Physician Health (2012 to 2016). We also searched the following association and foundation websites: American Medical Association, Australian Medical Association, British Medical Association, Canadian Medical Association, European Medical Association, National Sleep Foundation, Ontario Medical Association and the World Medical Association. The complete search strategy undertaken is reported in Supplementary file 1.

#### **Inclusion criteria**

Primary studies (quantitative or qualitative) of fatigue- or sleep-related exposures or interventions among physicians in independent practice were eligible for inclusion. We included physicians practicing in any medical specialty and in any healthcare setting within a high income country, [25] to identify practices comparable to the Canadian setting. Studies including physicians-in-training were included only if data for physicians in independent practice could be isolated. Exposures of interest included fatigue, sleep restriction, or sleepiness. We also included studies of any intervention that aimed to reduce fatigue or sleep restriction with any comparator (or no comparator). All reported outcomes, measured at any time, were eligible for inclusion.

We excluded commentaries, letters, editorials and dissertations. Systematic reviews, health technology assessments, economic evaluations and practice guidelines were excluded, although the reference lists of these as well as the included studies were scanned for potential primary studies. Studies that focused solely on physicians-in-training (e.g., trainees, residents, fellows, interns, medical students, junior doctors, registrars) were ineligible. To maintain the focused scope of the review, we excluded work hours, work load, and any other exposure or intervention that was indirectly related to fatigue or sleep restriction.

#### Study selection

The study team piloted the selection criteria, which were then applied by two independent reviewers following a two-phase process. We first screened titles and abstracts for potential relevance. Then, we retrieved all records classified as "include" or "unsure" and reviewed their full text for eligibility. Any

disagreements between reviewers were resolved by discussion or third-reviewer consultation when necessary.

#### **Data extraction**

Reviewers used a standardized form to extract data in Microsoft Office Excel (v. 2016, Microsoft Corporation, Redmond, WA). One reviewer independently extracted data from each included study and a second reviewer verified a random 10% sample. Since no major errors or omissions were noted, we did not undertake further verification.

We extracted the following data: country of publication; funding source; study design; inclusion and exclusion criteria; population characteristics (i.e., sample size, age and gender distribution, physician specialty); setting (i.e., physician workplace, urban or rural); exposure or intervention; definition of fatigue or sleep loss; sleep and fatigue scales used and timing of measurement; comparators (if applicable); and outcomes.

#### Risk of bias appraisal

Two reviewers independently assessed the risk of bias in each included study using standard tools. Disagreements were resolved via discussion or by consulting a third reviewer. We used the Cochrane Risk of Bias tool[22] to assess randomised controlled trials. Adapted versions of the tool developed by the Effective Practice and Organization of Care group[26] were used to assess before-after and time series studies. We used the Newcastle-Ottawa Quality Assessment Scale[27] to appraise cohort studies. We adapted the scale to assess cross-sectional studies and the one non-comparative study.

#### **Evidence synthesis**

We considered clinical and methodological heterogeneity in our decision on whether to proceed with meta-analysis for the outcomes identified. For most outcomes, we found insufficient homogeneity in study design, populations, exposures or interventions, and outcome measures to pool the data via meta-analysis. Thus, we have presented the findings for most outcomes narratively and in summary tables.

When statistical pooling was appropriate, this was undertaken using Review Manager (RevMan v.5.3, Copenhagen: The Nordic Cochrane Centre, the Cochrane Collaboration, 2014) via pairwise meta-analysis

When data were not presented in the format required for meta-analysis, we estimated means or standard deviations (SDs) using standard equations. We used the median instead of the mean for one study[31] for the outcomes of length of stay and operating time. Additionally, for one study[32] in the length of stay analysis where the SD could not be estimated, we substituted the mean variance of other studies within the meta-analysis.[33]

#### **RESULTS**

We identified 16,083 unique records via the database searches, 56 grey literature sources, and 14 additional records in reference lists of systematic reviews. We excluded 15,016 citations by title and abstract, and another 1,090 by full text. Forty-seven studies[31, 32, 34-78] were eligible for inclusion, and 6[31, 32, 41, 58, 63, 77] were included in meta-analysis for the outcomes of operating time, intra-and post-operative complications, patient mortality and length of hospital stay. Figure 1 shows the flow of studies through the selection process.

#### **Included study characteristics**

A summary of the study characteristics is provided in Table 1. Supplementary file 2 presents descriptive information for each included study. There were 45 observational studies [31, 32, 34-39, 41-74, 76-78] and two intervention studies. [40, 75] All studies were quantitative. Nearly half (n = 20/47, 43%) of the studies took place in North America, [31, 32, 35, 37, 38, 41, 45, 48, 57-60, 62, 63, 65-68, 72, 77] and slightly more than one-third (n = 16/47, 34%) in Europe. [34, 36, 39, 40, 42, 46, 47, 50-53, 61, 64, 73, 75, 76]

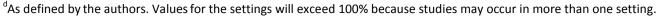
Table 1. Summary characteristics of the included studies

Study characteristics n %		%	Physician characteristics	n	%	Exposures, interventions and	n	%
						outcomes		
Study design			Gender		Exposures (observational) <sup>a</sup>	45	96	
Cross-sectional	34	72	Reported <sup>b</sup>	38	81	Fatigue-related	15	32
Cohort	6	13	>50% male	30	79	Sleep-related	37	79
Before-after	3	6	Age			Overnight or extended shifts	18	38
RCT	2	4	Reported <sup>b</sup>	38	81	Interventions (experimental)	2	4
Time series	1	2	Range (years)	20 to	>70	Outcomes		
Non-comparative	1	2	Specialty area <sup>c</sup> Physician health and wellbeing		Physician health and wellbeing	28	60	
Region and country			Surgeons	13	28	Work and life satisfaction	9	19
North America	20	43	Anesthesiologists	10	21	Burnout	7	15
US	15	32	Generalists	7	15	Stress	8	17
Canada	4	9	ED or ICU physicians	3	6	Mental health and wellbeing	7	15
Canada, US & Mexico	1	2	Oncologists	2	4	Other health-related outcomes	5	11
Europe	16	34	Obstetrician-gynecologists	1	2	Physician performance, risk of error	21	45
France	4	9	Mixed groups	14	30	Psychomotor performance	7	15
Finland	3	6	Work setting <sup>d</sup>		Work ability and quality of care	5	11	
Spain	2	4	Hospitals	37	79	Incidence of medical errors	5	11
Austria	2	4	Private practice	13	28	Surgical efficiency, effectiveness		13
Norway	2	4	Primary care centres, outpatient clinics	7	15	Patient outcomes		13
Denmark	1	2	Academic practice, training programs	5	11	1//		
Germany	1	2	Other (e.g., industry, military)	11	23			
Malta	1	2	Not reported 3 6					
Japan	4	9	Urban or rural					
Australia	2	4	Reported <sup>b</sup> 16 34					
Israel	2	4	Urban	12	75			
New Zealand	2	4	Rural	2	13			
United Kingdom	1	2	Mixed	2	13			

ED: emergency department; ICU: intensive care unit; RCT: randomised controlled trial; US: United States of America

<sup>&</sup>lt;sup>a</sup>Exposures that have been directly related to an outcome. Some studies included multiple exposures.

<sup>&</sup>lt;sup>c</sup>Anesthesiologists include physician anesthetists; generalists include primary care physicians, internists, and general practitioners; mixed groups refers to studies including more than one physician group or specialty (usually large-scale surveys). In some studies, multiple distinct groups were represented.





<sup>&</sup>lt;sup>b</sup>Percentages presented using the total number of studies where the outcome was reported as the denominator.

The 47 studies reported outcomes for 36,190 (range = 6 to 7,905) physicians and 69,809 (range = 270 to 38,978) adult patients. About half reported on surgeons (n = 13/47, 28%),[31, 32, 34, 41, 45, 48, 54, 58, 62, 63, 66, 75, 77] or anesthesiologists/physician anesthetists (n = 10/47, 21%).[36, 37, 39, 43, 44, 50-52, 59, 60] Where it was reported, the samples tended to be predominantly male (n = 30/38, 79%) and physician age varied widely. Hospitals were the most common setting (n = 37/47, 79%).[31, 32, 34-37, 39-43, 45-47, 49-52, 54, 57-61, 63-70, 73-75, 77, 78] In the studies where it was reported (n = 16/47, 34%),[31, 32, 34, 38, 40, 41, 43, 45, 50, 51, 55, 56, 58, 65, 76, 77] all but four studies[31, 55, 56, 77] took place in solely an urban setting.

Fifteen (32%) studies reported on fatigue-related exposures (e.g., as a source of stress, exhaustion, physical fatigue; hereafter referred to as 'fatigue'),[35, 40, 45, 48, 57, 63-71, 73] while others (n = 37/47, 79%) reported on sleep-related exposures (e.g., sleep hours, sleep restriction, sleep deprivation, sleep disruption, sleepiness; hereafter referred to as 'sleep restriction').[31, 32, 34, 36-47, 49-56, 58-62, 64, 67, 71, 72, 74-78] A few (n = 5/47, 11%) reported on both.[40, 45, 64, 67, 71] In some cases (n = 18/47, 38%), fatigue or sleep restriction were related to overnight work or long on-call shifts.[31, 32, 34, 37, 40, 41, 43, 45, 46, 50, 51, 53, 58-60, 63, 75, 77] Measured outcomes varied widely and were ultimately organised into physician physical and mental health, physician performance and risk of error, and patient outcomes.

#### Risk of bias appraisal

The overall quality of the body of research was poor; 62% (n = 29/47) of studies were rated at unclear or high risk of bias. Of the two randomised controlled trials, one was rated as unclear overall risk of bias[75] and one as high risk.[40] All cohort studies were at low risk of bias (mean score: 8.4/9, range: 8-9).[31, 32, 41, 58, 63, 77] All of the before-after studies were rated as high risk of bias.[34, 45, 50] The single time series study was assessed at high risk of bias.[51] The cross-sectional studies varied in performance (mean score: 3.0/5, range: 1-4); only one-third (n = 12/34, 35%) were at low risk of bias.[39, 42, 44, 47, 52, 59, 60, 69, 70, 72, 74, 76] The one non-comparative study was at unclear risk of bias.[43] Detailed assessments of the sources of bias per study are shown in Supplementary file 3.

### Physician health and wellbeing outcomes

Seven cross-sectional studies reported on burnout (5 low[39, 70, 72, 74, 76], 1 unclear[54], 1 high risk of bias[62]) among surgeons,[54, 62] anesthesiologists,[39] generalists,[76] and other mixed groups.[70, 72, 74] Two studies reported on surgeons; the larger (n = 2,564, low risk of bias) study of neurosurgeons showed increased odds of burnout with sleep deprivation (hours of sleep per night; OR 0.84, 95% CI 0.75 to 0.94, P = 0.002).[54] Among anesthesiologists one study (n = 565, low risk of bias) indicated that burnout (measured via Maslach Burnout Inventory) was more prevalent among the sleep-deprived ('lack of sleep' on one question; 47.6% vs. 16.3%, P < 0.001).[39] In one small (n = 11) study of generalists, those with burnout (measured via Pines Burnout Measure) had poorer Pittsburgh Sleep Quality Index scores (7.24±4.17 vs. 2.72±2.22, P < 0.001).[76] In the two larger studies of mixed physician groups (low risk of bias), burnout (measured via 5-point scale) was more prevalent among those who were sleep deprived (<7 hours of sleep per 24 hours; 39.6% vs. 26.4%, P < 0.05),[72] and physical fatigue ('feeling tired' on a 7-point scale) was correlated with burnout (Shirom-Melamed Burnout Measure; r = 0.88, P < 0.05).[70] In summary, evidence from 7 cross-sectional studies (71% at low risk of bias), showed associations between sleep restriction and burnout.

Six cross-sectional studies (2 low[47, 52], 1 unclear[46], 3 high risk of bias[35, 62, 64]), one uncontrolled before-after study (high risk of bias[50]), and one intervention study (high risk of bias[40]) reported on stress outcomes among surgeons,[62] anesthesiologists,[50, 52] emergency physicians,[40, 64] internal medicine physicians, [46] and mixed groups.[35, 47] In a small sample (n = 20) of internal medicine physicians, sleep restriction related to a 24-hour call shift showed no association with biochemical or physiological stress parameters, except levels of thyroid stimulating hormone, which was higher post-shift (P = 0.049, data not reported).[46] The remaining observational studies suggested that there was an association between sleep restriction or fatigue and stress. The one study of orthopedic surgeons (n = 264, high risk of bias) showed that sleep restriction (measured on a 3-point scale) and psychological distress (measured via General Health Questionnaire-12) were correlated (data not reported, P < 0.001).[62] The two reports on anesthesiologists were of varied quality; the larger (n = 328, low risk of bias) study showed that stress symptoms (measured via Modified Occupational Stress Questionnaire)

were predicted by sleep sufficiency (self-reported on one question,  $\beta$  = -0.269, P < 0.001).[52] Among the two studies reporting on mixed groups of physicians, the larger (n = 1,541, low risk of bias) study showed an association between sleep problems (4 questions derived from Jenkins scale) and psychological distress (General Health Questionnaire-12;  $\beta$  = 0.18, P < 0.001).[47] One RCT assessed the impact of sleep restriction from shift work (14-hour or 24-hour shifts), showing that stress (on a visual analog scale) among emergency physicians (n = 17) was higher following the shift as compared to a control day (data not reported, P < 0.05).[40] In summary, evidence from one intervention study at high risk of bias and all but one of the 7 observational studies (29% at low risk of bias) identified supported an inverse association between fatigue or sleep deprivation and stress.

Seven cross-sectional studies (2 low, [52, 60] 3 unclear, [67, 71, 78] 2 high risk of bias [36, 53]) reported on aspects of mental health including addiction or substance misuse, [36, 53, 71] depression, [78] thoughts of suicide, [52] mood disturbance, [60, 71] and overall wellbeing. [67] One study, [53] which was at high risk of bias, showed no association between hours of sleep when on call and hazardous drinking behaviours (via Alcohol Use Disorder Identification Test). Meanwhile, the six other studies all showed associations between sleep restriction and fatigue and reduced mental health. Three studies reported on anesthetists, [36, 52, 60] with two large surveys showing increased odds of tobacco (OR 1.42, 95% CI 1.04 to 1.94) and tranquilizer/hypnotics (OR 3.26, 95% CI 2.12 to 5.02) dependency being predicted by sleep deprivation (measured by one question),[36] and sleep disturbance being associated with thoughts of suicide (using a 4-point scale; P = 0.009).[52] A small study (n = 21) showed greater mood disturbance following a 17-hour night shift than a usual day (Profile of Mood States score 42.57±15.26 vs. 70.90±6.91, P < 0.001).[60] Among oncologists (n = 241), overall wellbeing was predicted by lower levels of fatigue after controlling for personal and professional characteristics (assessed via linear analog scale quality of life survey, P = 0.002).[67] A large (n = 3,862, unclear risk of bias) study of physicians showed that sleep restriction (lower sleep hours when not at work in the past month) was associated with increased odds of depression (Quick Inventory Depressive Scale; OR 2.70, 95% CI 1.82 to 4.03 for men; OR 2.38, 95% CI 1.11 to 5.10 for women).[78] In open-ended questions, senior physicians in one study (unclear risk of bias) attributed the development of mental illness to tiredness and stress at work.[71] In summary, 7 cross-sectional studies (29% at low risk of bias) were identified, and of these 6 supported an association between sleep restriction or fatigue and negative mental health outcomes.

Nine cross-sectional studies (4 low,[42, 47, 72, 74] 2 unclear,[55, 68] 3 high risk of bias[38, 48, 62]) reported on outcomes related to job satisfaction,[42, 47, 48, 55, 72, 74] life satisfaction,[38, 62, 72] or work-life balance.[68, 72] The six studies that investigated job satisfaction were all at low risk of bias and generally included mixed groups of physicians;[47, 72, 74] one study reported on general practitioners,[55], another on surgeons,[48] and one on mixed specialties.[42] Three studies showed that reductions in sleep duration and/or quality[47, 48, 74] were associated with reduced job satisfaction. Meanwhile one showed no association between sleep restriction (<7 hours per 24-hour period) and career satisfaction (measured on a 5-point Likert scale),[72] and another showed no relationship between earlier sleep disturbance (Jenkins Scale) and later job demands or job control (measured via 5-point scale).[42] A single study (n = 92) reporting on rural general practitioners indicated that frequent sleep disturbance (measured on a 7-point scale) predicted the intention to retire early (OR 2.91, 95% CI 1.11 to 7.6, P < 0.05).[55] In summary, 6 cross-sectional studies (all at low risk of bias) were identified, and all but two[42, 72] of these studies showed that sleep restriction and fatigue were associated with reductions in satisfaction.

The three studies reported on life satisfaction. [38, 62, 72] Of two studies among mixed physician groups, [38, 72] the one larger (n = 840) study showed that sleep restriction (< 7 hours per day) was a predictor of reduced life satisfaction (measured on a 5-point Likert scale; OR 0.44, 95% CI 0.29 to 0.67, P  $\leq$  0.05). [72] One study at high risk of bias reported on orthopedic surgeons (n = 264), showing that sleep deprivation (measured via 3-point scale) was correlated with lower marital satisfaction (Revised Dyadic Adjustment Scale; data not reported, P < 0.001). [62] Two large studies at low or unclear risk of bias reported on work-life balance. [68, 72] Among oncologists (n = 1,117), reduced satisfaction with work-life balance (measured on a 5-point Likert scale) was predicted by high levels of fatigue (measured via 10-point visual analog scale), even when controlling for personal and work-related factors and burnout (OR 0.489, 95% CI 0.337 to 0.710, P < 0.001). [68] Among a mixed group of physicians (n = 840, low risk of bias), sleep restriction (<7 hours in a typical 24-hour period) predicted a reduced perception of having balanced personal and professional commitments (5-point Likert scale; OR 0.46, 95% CI 0.31 to 0.71, P  $\leq$  0.05). [72] In summary, 3 cross-sectional studies (all unclear or high risk of bias) supported an association between sleep restriction or fatigue and reduced life satisfaction, and 2 cross-sectional studies (50% low risk of bias) supported an association with reduced work-life balance.

Four cross sectional studies (3 unclear,[56, 57, 71] 1 high risk of bias[38]) and one time series study (high risk of bias[51]) reported on other health-related outcomes. Among a mixed group of physicians (n = 180), one study at high risk of bias showed that Epworth Sleepiness Scale scores were higher among physicians who worried about having a car accident while driving home (7.0 vs. 5.4, P < 0.001).[38] Among generalists (n = 578), almost 1 in 10 (8.7%) admitted to falling asleep while driving due to fatigue.[57] Also among generalists (n = 92), those with frequent work-related sleep disturbance (measured on a 7-point scale) were at increased odds of sickness presenteeism (OR 2.92, 95% CI 1.19 to 7.16, P = 0.02).[56] The one time series study concluded that a single 24-h shift did not cause major chronodisruption (based on serum melatonin measurement) among anesthetists (n = 10).[51] Meanwhile, open-ended comments from a large sample (n = 3,550) of senior physicians suggests that they attributed the development of physical health problems to a lifestyle of sleep restriction, poor eating habits and lack of exercise imposed by their jobs.[71] In summary, 5 cross sectional studies (0% at low risk of bias) supported associations between sleep restriction and fatigue and varied deleterious health outcomes (i.e., car accidents, sickness presenteeism, physical health problems). One time series study at high risk of bias did not support such a relationship.

#### Physician performance and risk of errors

Twenty-one studies reported on physician performance and safety-related outcomes, [31, 32, 34, 37, 38, 41, 43-47, 49, 50, 59, 61, 65, 66, 69, 71, 73, 75] including surgical efficiency and effectiveness (n = 6), psychomotor performance (n = 7), work ability and quality of care (n = 5) and medical errors (n = 5) (Supplementary file 4).

Four cohort studies (all low risk of bias[31, 32, 41, 63]), one before-after study (high risk of bias[34]) and one randomized controlled trial (high risk of bias[75]) examined the effects of sleep restriction from overnight work or extended shifts, during surgeries[31, 32, 41] or laparoscopic simulations.[34, 75] We pooled the data from these studies[31, 32, 41, 63] via meta-analysis, which showed no difference in operating time (sometimes referred to as surgeon efficiency) between sleep deprived and non-sleep deprived surgeons (Figure 2; n = 50,046, MD -0.14, 95% CI -1.60 to 1.33, P = 0.86,  $I^2 = 0\%$ ). Of studies not meta-analysed, the small (n = 29) before-after study showed no impact of sleep deprivation from shiftwork nor of sleep hours on performance on a laparoscopic simulation (LapSimGyn).[34] One small (n = 64) intervention study compared a 24-hour shift to a usual work day, also finding no detriment to performance on a laparoscopic simulation (Minimally Invasive Surgical Trainer-Virtual Reality) despite

diminished sleep hours while working on-call.[75] In summary, pooled data from 4 cohort studies (100% low risk of bias) showed no effect of sleep restriction on surgical efficiency. Additional data from one RCT (high risk of bias) and one before-after study (high risk of bias) also showed no association between sleep restriction and performance on laparoscopic simulations.

Two before-after studies (high risk of bias[45, 50]) and five cross-sectional studies (2 low,[43, 59] 3 unclear, [37, 46] 1 high risk of bias [61]) reported on psychomotor performance outcomes among surgeons,[45] anesthesiologists,[37, 43, 50, 59] emergency physicians,[61] and internal medicine physicians.[46] Among a small group of surgeons (n = 9), performance on a virtual ring transfer task deteriorated after an on-call shift (data not reported, P < 0.05).[45] The four studies among anesthetists reported mixed findings. One small (n = 11) before-after study showed longer reaction times (690.8±73.4 vs. 746.5±113.7 milliseconds) and reduced concentration ability (26.4±23.5 vs. 56.3±23.0 on a 100-point scale, P = 0.007) following a 24-hour shift with sleep restriction; [50] Two others found that sleep restriction due to overnight shifts was associated with slower reaction times. [43, 59] Conversely, a small study (n = 11) found no effect of overnight shiftwork with sleep restriction on any measure of psychomotor performance except Hopkin's Verbal Learning Test (t-score of 48.6±7.6 vs. 41.5±9.9, P = 0.04).[37] Among emergency physicians (n = 18), one study (high risk of bias) showed that those who were sleep deprived (<5 hours sleep after a 24-hour shift) had a reduced performance on most but not all psychomotor tests (Battery Test Reaction 5),[61] while among internal medicine physicians (n = 20, low risk of bias), neurocognitive parameters did not seem to worsen post-call.[46] In summary, two before-after (0% low risk of bias) and 5 cross-sectional studies (40% low risk of bias) showed mixed results for the association between fatigue or sleep restriction and psychomotor performance.

Five cross-sectional studies (2 low,[47, 69] 1 unclear,[71] 2 high risk of bias[38, 65]) reported on associations between sleep deprivation or fatigue and work ability or perceived performance, all among mixed groups of physicians.[38, 47, 65, 69, 71] The two large studies at low risk of bias showed that sleep problems and fatigue were inversely associated with physicians' perceived quality of work.[47, 69] Among 1,541 physicians in Finland, sleeping problems (measured by 4 questions from the Jenkins Scale) were inversely associated with scores on the Work Ability Index ( $\beta$  = -0.29, P < 0.001),[47] while a study of 890 physicians from Israel demonstrated that perceived quality of care was predicted by fatigue (1 item on the Shirom-Melamed Burnout Measure) even after controlling for components of burnout ( $\beta$  = 0.17, P < 0.05).[69] Similarly, in one study, comments from senior physicians suggested that continual

tiredness and exhaustion negatively affected their perceived competence. [71] The two studies [38, 65] that were at high risk of bias had conflicting findings. In summary, 5 cross-sectional studies (40% at low risk of bias) reported on perceived work performance; those that were at low risk of bias supported an association between fatigue or sleep restriction and reduced performance.

Five cross-sectional studies (1 low,[44] 2 unclear,[49, 66] 2 high risk of bias[38, 73]) reported on associations between sleep restriction or fatigue and self-reported medical errors among surgeons,[66] anesthesiologists[44] and mixed groups of physicians.[38, 49, 73] A large (n = 7,905) study at unclear risk of bias showed that only 6.9% of surgeons reported fatigue as the most important contributor to medical errors.[66] Among anesthesiologists, a smaller study (n = 183) at low risk of bias showed that the risk of self-reported fatigue-related errors increased with more nights of work-related sleep disturbance (RR 1.25, 95% CI 1.06 to 1.49).[44] Two of the studies reporting on mixed groups of physicians had conflicting results,[38, 49] while another reported that physicians' opinions on the association between fatigue and prescribing errors differed by work setting.[73] One-third (34%) of community-based, 96% of hospital-based, and 8% of office-based physicians believed that there was a high or very high association between fatigue and prescribing errors (P < 0.05).[73] In summary, 5 cross-sectional studies (20% at low risk of bias) reported on self-reported errors, and these showed mixed findings for associations with fatigue or sleep restriction.

#### **Patient Outcomes**

Six large (n = 270 to 38,978) cohort studies at low risk of bias reported on patient outcomes, all related to surgical[31, 32, 41, 58, 63, 77] or obstetric[58] procedures (Supplementary file 4). In these studies, sleep restriction or fatigue were typically defined as overnight work prior to a daytime procedure[31, 41, 58, 63, 77]; though two studies measured sleep hours[32] or 'sleep opportunity'.[58] We pooled data for procedures performed by sleep deprived versus non-sleep deprived surgeons (or obstetrician-gynecologists in one case[58]). Analyses showed no difference in the rate of intra-operative complications (Figure 3, 3 studies,[58, 63, 77] n = 19,798, RR 1.35, 95% CI 0.82 to 2.21, p=0.24,  $I^2 = 82\%$ ), post-operative complications (Figure 4; 5 studies,[31, 32, 41, 63, 77] n = 60,201, RR 0.99, 95% CI 0.95 to 1.03, p = 0.51,  $I^2 = 0\%$ ), patient mortality (Figure 5; 5 studies,[31, 32, 41, 63, 77] n = 60,436, RR 0.98, 95% CI 0.84 to 1.15, p = 0.82,  $I^2 = 0\%$ ), or length of hospital stay in days (Figure 6; 4 studies,[31, 32, 41, 63] n = 50,046, MD -0.33, 95% CI -1.03 to 0.36, p = 0.35,  $I^2 = 86\%$ ). One study[77] in the mortality analysis reported the number of deaths only as  $\leq 5$ . We assumed 2 events for this study (midpoint between 0 and

5); sensitivity analysis using the lowest (i.e., 0) and highest (i.e., 5) possible number of events did not change the overall result (Supplementary file 5). We imputed the average variance for one study[32] in the length of stay analysis; sensitivity analysis using either the highest or lowest SD did not change the results (Supplementary file 5). Subgroup analysis by type of surgery did not explain the substantial between-study heterogeneity detected for length of stay, nor intraoperative complications, though it may be noted that the types of complications reported varied by study.

#### **DISCUSSION**

Fatigue and chronic sleep restriction are two potential drivers of reduced physician wellbeing[17, 19] that have thus far been understudied in physicians in independent practice. Burnout is becoming increasingly prevalent among physicians,[14-16] and recent research indicates that comprehensive individual- and system-level strategies are needed to address the problem.[6-9, 19, 21] We have systematically reviewed evidence from a heterogeneous array of available studies reporting on diverse outcomes related to physicians in independent practice and their patients. The included studies were often at high or unclear risk of bias, included small samples of physicians, and inconsistently measured and reported exposures and outcomes. The key message gleaned from this review is that despite growing interest in the topic of physician wellness, the robust evidence needed to inform individual and systems-level fatigue management strategies is lacking.

Traditionally, much of the fatigue-related research has focused on hazards to patients. The current review included six cohort studies showing that sleep restriction and/or fatigue did not seem to result in increased rates of patient morality, operative complications, or length of hospital stay. Despite these findings, evidence for psychomotor performance, surgical skills and errors suggest that there is indeed a potential for negative outcomes. The included studies, like many of the others in this and other systematic reviews,[79] employed indirect definitions that make it difficult to classify sleep deprived physicians with certainty. In recent years there has been a shift away from the singular focus on patient safety toward a more comprehensive view that also considers the detrimental effects of fatigue, sleep loss and other occupational hazards on physician wellness.[80] Evidence from this review supports that fatigue and sleep restriction may be negatively associated with physician health and wellbeing. It is now recognized that health systems cannot be sustained by a workforce that is facing an epidemic of burnout.[19, 81, 82]

In light of high rates of burnout, the ongoing dialogue about the need for a cultural shift in the practice of medicine[83, 84] is now more important than ever. Recognition of the potential effects of physician fatigue on patients, physicians, and healthcare systems as a whole must be emphasized at a systemic level, encouraging a shift in which the risks are viewed as unacceptable.[1, 20, 80] Likewise, although research to date has focused largely on individual-level approaches to address burnout, it is now clear that placing the burden of a system-level problem solely on the individual is unlikely to bring about significant and lasting change.[85] Recent research has highlighted physician burnout as a system-driven issue that will require corresponding national-scale multicomponent solutions.[1, 19, 81, 82] As such, in the past several years both the American and Canadian Medical Associations have developed policies and programs that address physician health.[81, 86] The Canadian Medical Association's new policy on physician health calls on broad stakeholder groups (e.g., policymakers, regional health authorities, governments) to take shared responsibility for the health of physicians and to make meaningful and concerted efforts towards promoting a healthy and sustainable workforce.[81]

The most salient finding of this review is that the current evidence is insufficient to inform policy and practice. Correspondingly, a 2016 research summit on physician wellness and burnout outlined the need for timely, relevant and methodologically robust research to inform practice and policy.[21] The findings herein may be used by researchers and practitioners to develop and design methodologically strong research programs related to physician fatigue, inform successful research grant proposals, and lobby healthcare organizations to increase the focus on physician fatigue management programs. It will be important to make use of existing validated measures[87-89] consistently in future research. Identifying outcomes of importance to physicians and their patients should be prioritized, such that these may be collected within intervention studies. Reporting these consistently will allow for the effective synthesis of findings and reduce research waste.[90] Integrated knowledge translation strategies involving multiple stakeholder groups (e.g., physicians, patients, medical schools, physicians' associations and governing bodies, policymakers) may help to ensure that the research is relevant and facilitates decision-making.[91]

#### **Strengths and Limitations**

Our systematic review is the first to synthesize evidence on the effects of fatigue and sleep restriction on physicians in independent practice. The review is timely, given recent calls for research into individual and organisational solutions for burnout, [20, 21] and an increased focus on physician health. [80, 81]

some detrimental physician health and wellbeing outcomes; the evidence for potential associations with performance and safety outcomes was mixed. Meta-analyses for patient outcomes did not show any significant associations with physician sleep deprivation. Our overall confidence in the findings is low, owing to a body of research that is hindered by methodological weaknesses. Further methodologically robust research that includes consistent outcomes that are of interest to physicians and their patients is needed to inform strong practice recommendations and policy decisions.

#### **ACKNOWLEDGMENTS**

We would like to thank Dr. Meghan Sebastianski and Andrea Milne for assistance with screening, and Dr. Allison Gates for screening, data verification, and quality assessment of the French studies.

#### COPYRIGHT/LICENSE FOR PUBLICATION

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide license to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material whereever it may be located; and, vi) license any third party to do any or all of the above.

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

#### **COMPETING INTERESTS**

All authors have completed the ICMJE uniform disclosure form at <a href="www.icmje.org/coi\_disclosure.pdf">www.icmje.org/coi\_disclosure.pdf</a> and declare: Dr. Christopher Simon is employed by the Canadian Medical Association, who provided financial support for the research; there are no other relationships or activities that could appear to have influenced the submitted work.

#### **CONTRIBUTOR STATEMENT**

All authors contributed to the conception and design of the project. MG and AW contributed to the acquisition, analysis and interpretation of the data, and drafted the manuscript. RF contributed to acquisition of data. CSa, CSi and MPD contributed to interpretation of data and revised the manuscript for important intellectual content. All authors approved the final version of the manuscript as submitted.

#### **FUNDING STATEMENT**

This research was supported by the Canadian Medical Association. The funder had no role in the study design; collection, analysis and interpretation of data; the writing of the report; and the decision to submit the article for publication.

#### **ROLE OF FUNDERS**

Dr. Christopher Simon is employed by the Canadian Medical Association. The remaining authors are independent from the funders. The funders had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

#### TRANSPARENCY DECLARATION

The lead author (MG) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; no important aspects of the study have been omitted; and all discrepancies from the study as planned have been explained.

#### **DATA ACCESS STATEMENT**

All authors, external and internal, had full access to all of the data in the study and can take responsibility for the integrity of the data and the accuracy of the interpretation.

# BMJ Open: first published as 10.1136/bmjopen-2018-021967 on 21 September 2018. Downloaded from http://bmjopen.bmj.com/ on June 9, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies. **DATA SHARING STATEMENT** The data pertaining to this systematic review are available from the corresponding author upon reasonable request. FIGURE CAPTIONS **Figure 1.** Flow of records through the selection process Figure 2. Forest plot for operating time among sleep deprived and non-sleep deprived surgeons Figure 3. Forest plot for intra-operative complications among procedures performed by sleep deprived and non-sleep deprived surgeons or obstetrician-gynecologists Legend: Rothschild 2009 reported the total number of procedures with complications; Schieman 2008 reported the intraoperative complication rate; Vinden 2013 reported conversion to open procedure Figure 4. Forest plot for post-operative complications among surgeries performed by sleep deprived and non-sleep deprived surgeons Legend: Vinden 2013 reported iatrogenic injuries; Schieman 2008, Govindarajan 2015, and Chu 2011 reported post-operative complication rate; Ellman 2004 reported post-operative complications (other types of complications reported not included in the analysis) Figure 5. Forest plot for patient mortality among surgeries performed by sleep deprived and non-sleep deprived surgeons Figure 6. Forest plot for patient length of hospital stay (days) among surgeries performed by sleep deprived and non-sleep deprived surgeons

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text and

#### REFERENCES

- 1. Greig P, Snow R. Fatigue and risk: are train drivers safer than doctors? *BMJ* 2017;359. doi: https://doi.org/10.1136/bmj.j5107
- 2. Parker JBR. The effects of fatigue on physician performance an underestimated cause of physician impairment and increased patient risk. *Can J Anaesth* 1987;34(5):489-95.
- 3. Temple J. Resident duty hours around the globe: where are we now? BMC Med Educ 2014;14(Suppl 1):S8. https://doi.org/10.1186/1472-6920-14-S1-S8
- 4. Imrie K, Frank J, The National Steering Committee on Resient Duty Hours. Fatigue, risk, & excellence: towards a pan-Canadian consensus on resident duty hours. Ottawa, Ontario: The Royal College of Physicians and Surgeons of Canada, 2013. <a href="http://www.residentdutyhours.ca/final\_report.php">http://www.residentdutyhours.ca/final\_report.php</a> (accessed 12 Jan 2018).
- Accreditation Council for Graduate Medical Education. History of duty hours. 2017.
   <a href="http://www.acgme.org/What-We-Do/Accreditation/Clinical-Experience-and-Education-formerly-Duty-Hours/History-of-Duty-Hours">http://www.acgme.org/What-We-Do/Accreditation/Clinical-Experience-and-Education-formerly-Duty-Hours/History-of-Duty-Hours</a> (accessed 12 Jan 2018).
- 6. Bolster L, Rourke L. The effect of restricting residents' duty hours on patient safety, resident well-being, and resident education: an updated systematic review. *J Grad Med Educ* 2015;7(3):349-63.
- 7. Harris JD, Staheli G, LeClere L, Andersone D, McCormick F. What effects have resident work-hour changes had on education, quality of life, and safety? A systematic review. *Clin Orthop* 2015;473(5):1600-8.
- 8. Ahmed N, Devitt KS, Keshet I, et al. A systematic review of the effects of resident duty hour restrictions in surgery: impact on resident wellness, training, and patient outcomes. *Ann Surg* 2014;259(6):1041-53.
- 9. Peets A, Ayas NT. Restricting resident work hours: the good, the bad, and the ugly. *Crit Care Med* 2012;40(3):960-6.
- 10. Canadian Medical Association. CMA Policy: Management of physician fatigue. Ottawa, Canada: Canadian Medical Association; 2014.
- 11. Anim M, Markert RJ, Wood VC, Schuster BL. Physician practice patterns resemble ACGME duty hours. *Am J Med* 2009;122:587-93.
- 12. Transport Canada. Fatigue risk management system for the canadian aviation industry: policies and procedures development guidelines. Ottawa, Canada: Her Majesty the Queen in Right of Canada, as represented by the Minister of Transport, 2008.
  - https://www.tc.gc.ca/eng/civilaviation/publications/TP14576-6042.htm (accessed 12 Jan 2018).

14. Boudreau RA, Grieco RL, Cahoon SL, Robertson RC, Wedel RJ. The pandemic from within: two surveys of physician burnout in Canada. *Can J Commun Ment Health* 2006;25(2):71-88.

- 15. Kumar S. Burnout and doctors: prevalence, prevention and intervention. *Healthcare* 2016;4(3):37. doi: 10.3390/healthcare4030037
- 16. Shanafelt TD, Boone S, Tan L, et al. Burnout and satisfaction with work-life balance among us physicians relative to the general us population. *Arch Intern Med* 2012;172(18):1377-85.
- 17. Brady KJS, Trockel MT, Khan CT, et al. What do we mean by physician wellness? A systematic review of its definition and measurement. *Acad Psychiatry* 2017. doi: 10.1007/s40596-017-0781-6
- 18. Fralick M, Flegel K. Physician burnout: who will protect us from ourselves? *CMAJ* 2014;186(10):731. doi: 10.1503/cmaj.140588
- 19. Shanafelt TD, Noseworthy JH. Executive leadership and physician well-being: nine organizational strategies to promote engagement and reduce burnout. *Mayo Clin Proc* 2017;92(1):129-46.
- 20. Lemaire JB, Wallace JE. Burnout among doctors. BMJ 2017;358:j3360. doi: 10.1136/bmj.j3360
- 21. Dyrbye LN, Trockel M, Frank E, et al. Development of a research agenda to identify evidence-based strategies to improve physician wellness and reduce burnout. *Ann Intern Med* 2017;166(10):743-4.
- 22. Higgins JPT, Green S (editors). The Cochrane handbook for systematic reviews of interventions, version 5.1.0. London, UK: The Cochrane Collaboration, 2011. <a href="http://www.handbook.cochrane.org">http://www.handbook.cochrane.org</a> (accessed 12 Jan 2018).
- 23. 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(7):e1000100. https://doi.org/10.1371/journal.pmed.1000100
- 24. Pattani R, Wu PE, Dhalla IA. Resident duty hours in Canada: past, present and future. *CMAJ* 2014;186(10):761-5.
- 25. International Monetary Fund. World economic outlook: too slow for too long. Washington, DC: International Monetary Fund, 2016. <a href="http://www.imf.org/external/pubs/ft/weo/2016/01/">http://www.imf.org/external/pubs/ft/weo/2016/01/</a> (accessed 12 Jan 2018).
- 26. Cochrane Effective Practice and Organisation of Care. Suggested risk of bias criteria for EPOC reviews. Oslo, Norway: Norwegian Knowledge Centre for the Health Services, 2016. http://epoc.cochrane.org/epoc-specific-resources-review-authors (accessed 12 Jan 2018).

- 27. Wells GA SB, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa, Canada: The Ottawa Hospital Research Institute, 2014.

  http://www.ohri.ca/programs/clinical\_epidemiology/oxford.asp (accessed 12 Jan 2018).
- 28. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7(3):177-88.
- 29. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-60.
- 30. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315(7109):629-34.
- 31. Govindarajan A, Urbach DR, Kumar M, et al. Outcomes of daytime procedures performed by attending surgeons after night work. *N Engl J Med* 2015;373(9):845-53.
- 32. Chu MW, Stitt LW, Fox SA, Kiaii B, Quantz M, Guo L, et al. Prospective evaluation of consultant surgeon sleep deprivation and outcomes in more than 4000 consecutive cardiac surgical procedures. *Arch Surg* 2011;146(9):1080-5.
- 33. Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *J Clin Epidemiol* 59(1):7-10.
- 34. Amirian I, Andersen LT, Rosenberg J, Gogenur I. Laparoscopic skills and cognitive function are not affected in surgeons during a night shift. *J Surg Educ* 2014;71(4):543-50.
- 35. Aziz A. Sources of perceived stress among american medical doctors: a cross-cultural perspective. Cross Cultural Management 2004;11(4):28-39.
- 36. Beaujouan L, Czernichow S, Pourriat JL, Bonnet F. Prevalence and risk factors for substance abuse and dependence among anaesthetists: a national survey. *Ann Fr Anesth Réanim* 2005;24(5):471-9.
- 37. Chang LC, Mahoney JJ, 3rd, Raty SR, Ortiz J, Apodaca S, De La Garza R 2nd. Neurocognitive effects following an overnight call shift on faculty anesthesiologists. *Acta Anaesthesiol Scand* 2013;57(8):1051-7.
- 38. Chen I, Vorona R, Chiu R, Ware JC. A survey of subjective sleepiness and consequences in attending physicians. *Behav Sleep Med* 2008;6(1):1-15.
- 39. Doppia MA, Estryn-Béhar M, Fry C, Guetarni K, Lieutaud T, comité de pilotage de l'enquête SESMAT. Burnout in French doctors: a comparative study among anaesthesiologists and other specialists in French hospitals (SESMAT study). *Ann Fr Anesth Réanim* 2011;30(11):782-94.

- 41. Ellman PI, Law MG, Tache-Leon C, et al. Sleep deprivation does not affect operative results in cardiac surgery. *Ann Thorac Surg* 2004;78(3):906-11.
- 42. Elovainio M, Heponiemi T, Jokela M, et al. Stressful work environment and wellbeing: what comes first? *J Occup Health Psychol* 2015;20(3):289-300.
- 43. Gander P, Millar M, Webster C, Merry A. Sleep loss and performance of anaesthesia trainees and specialists. *Chronobiol Int* 2008;25(6):1077-91.
- 44. Gander PH, Merry A, Millar MM, Weller J. Hours of work and fatigue-related error: a survey of New Zealand anaesthetists. *Anaesth Intensive Care* 2000;28(2):178-83.
- 45. 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(6):813-9.
- 46. Harbeck B, Suefke S, Haas CS, Lehnert H, Kropp P, Moenig H. No stress after 24-hour on-call shifts? *J Occup Health* 2015;57(5):438-47.
- 47. Heponiemi T, Puttonen S, Elovainio M. On-call work and physicians' well-being: testing the potential mediators. *Occup Med* 2014;64(5):352-7.
- 48. Jackson TN, Pearcy CP, Khorgami Z, Agrawal V, Taubman KE, Truitt MS. The physician attrition crisis: a cross-sectional survey of the risk factors for reduced job satisfaction among US surgeons. *World J Surg* 2017;24. doi: 10.1007/s00268-017-4286-y
- 49. Kaneita Y, Ohida T. Association of current work and sleep situations with excessive daytime sleepiness and medical incidents among Japanese physicians. *J Clin Sleep Med* 2011;7(5):512-22.
- 50. Lederer W, Kopp M, Hahn O, et al. Post-duty psychomotor performance in young and senior anaesthetists. *Eur J Anaesthesiol* 2006;23(3):251-6.
- 51. Leichtfried V, Putzer G, Perkhofer D, Schobersberger W, Benzer A. Circadian melatonin profiles during single 24-h shifts in anesthetists. *Sleep Breath* 2011;15(3):503-12.
- 52. Lindfors PM, Nurmi KE, Meretoja OA, et al. On-call stress among Finnish anaesthetists. *Anaesthesia* 2006;61(9):856-66.
- 53. Mahmood JI, Stoen Grotmol K, Tesli M, Vaglum P, Tyssen R. Contextual factors and mental distress as possible predictors of hazardous drinking in Norwegian medical doctors: a 15-year longitudinal, nationwide study. *Eur Addict Res* 2017;23(1):19-27.

- 54. Nishimura K, Nakamura F, Takegami M, et al. Cross-sectional survey of workload and burnout among Japanese physicians working in stroke care: the nationwide survey of acute stroke care capacity for proper designation of comprehensive stroke center in Japan (J-ASPECT) study. *Circ Cardiovasc Qual Outcomes* 2014;7(3):414-22.
- 55. Pit SW, Hansen V. Factors influencing early retirement intentions in Australian rural general practitioners. *Occup Med* 2014;64(4):297-304.
- 56. Pit SW, Hansen V. The relationship between lifestyle, occupational health, and work-related factors with presenteeism amongst general practitioners. *Arch Environ Occup Health* 2016;71(1):49-56.
- 57. Roberts DL, Shanafelt TD, Dyrbye LN, West CP. A national comparison of burnout and work-life balance among internal medicine hospitalists and outpatient general internists. *J Hosp Med* 2014;9(3):176-81.
- 58. Rothschild JM, Keohane CA, Rogers S, et al. Risks of complications by attending physicians after performing nighttime procedures. *JAMA* 2009;302(14):1565-72.
- 59. Saadat H, Bissonnette B, Tumin D, et al. Effects of partial sleep deprivation on reaction time in anesthesiologists. *Paediatr Anaesth* 2017;27(4):358-62.
- 60. Saadat H, Bissonnette B, Tumin D, et al. Time to talk about work-hour impact on anesthesiologists: The effects of sleep deprivation on Profile of Mood States and cognitive tasks. *Paediatr Anaesth* 2016;26(1):66-71.
- 61. Sanches I, Teixeira F, dos Santos JM, Ferreira AJ. Effects of acute sleep deprivation resulting from night shift work on young doctors. *Acta Med Port* 2015;28(4):457-62.
- 62. Sargent MC, Sotile W, Sotile MO, Rubash H, Barrack RL. Quality of life during orthopaedic training and academic practice. Part 1: orthopaedic surgery residents and faculty. *J Bone Joint Surg Am* 2009;91(10):2395-405.
- 63. Schieman C, MacLean AR, Buie WD, Rudmik LR, Ghali WA, Dixon E. Does surgeon fatigue influence outcomes after anterior resection for rectal cancer? *Am J Surg* 2008;195(5):684-7.
- 64. Sende J, Jbeili C, Schvahn S, Khalid M, Asaph J, Romano H, et al. Stress factors and stress consequences among emergency physicians: vational survey. *Annales Françaises de Médecine d'Urgence* 2012;2(4):224-31.
- 65. Sexton JB, Thomas EJ, Helmreich RL. Error, stress, and teamwork in medicine and aviation: Cross sectional surveys. *Hum Perform Extrem Environ* 2001;6(1):6-11.
- 66. Shanafelt TD, Balch CM, Bechamps G, et al. Burnout and medical errors among American surgeons. *Ann Surg* 2010;251(6):995-1000.

- 68. Shanafelt TD, Raymond M, Kosty M, et al. Satisfaction with work-life balance and the career and retirement plans of US oncologists. *J Clin Oncol* 2014;32(11):1127-35.
- 69. Shirom A, Nirel N, Vinokur AD. Overload, autonomy, and burnout as predictors of physicians' quality of care. *J Occup Health Psychol* 2006;11(4):328-42.
- 70. Shirom A, Nirel N, Vinokur AD. Work hours and caseload as predictors of physician burnout: the mediating effects by perceived workload and by autonomy. *J Appl Psychol* 2010;59(4):539-65.
- 71. Smith F, Goldacre MJ, Lambert TW. Adverse effects on health and wellbeing of working as a doctor: views of the UK medical graduates of 1974 and 1977 surveyed in 2014. *J R Soc Med* 2017;110(5):198-207.
- 72. Starmer AJ, Frintner MP, Freed GL. Work-life balance, burnout, and satisfaction of early career pediatricians. *Pediatrics* 2016;e20153183. doi: 10.1542/peds.2015-3183
- 73. Tanti A, Camilleri M, Borg AA, et al. Opinions of Maltese doctors and pharmacists on medication errors. *Int J Saf Med* 2017;29(1-2):81-99.
- 74. Tokuda Y, Hayano K, Ozaki M, Bito S, Yanai H, Koizumi S. The interrelationships between working conditions, job satisfaction, burnout and mental health among hospital physicians in Japan: a path analysis. *Ind Health* 2009;47(2):166-72.
- 75. 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(6):753-7.
- 76. Vela-Bueno A, Moreno-Jimenez B, Rodriguez-Munoz A, et al. Insomnia and sleep quality among primary care physicians with low and high burnout levels. *J Psychosom Res* 2008;64(4):435-42.
- 77. Vinden C, Nash DM, Rangrej J, et al. Complications of daytime elective laparoscopic cholecystectomies performed by surgeons who operated the night before. *Obstet Gynecol Surv* 2014;69(2):71-3.
- 78. Wada K, Yoshikawa T, Goto T, et al. National survey of the association of depressive symptoms with the number of off duty and on-call, and sleep hours among physicians working in Japanese hospitals: a cross sectional study. *BMC Public Health* 2010;10:127. doi: 10.1186/1471-2458-10-127
- 79. Sturm L, Dawson D, Vaughan R, et al. Effects of fatigue on surgeon performance and surgical outcomes: a systematic review. *ANZ J Surg* 2011;81(7-8):502-9.

- 80. Wallace JE, Lemaire JB, Ghali WA. Physician wellness: a missing quality indicator. *Lancet* 2009;374(9702):1714-21.
- 81. Canadian Medical Association. CMA policy: physician health. Ottawa, Canada: Canadian Medical Assocation, 2017. <a href="http://policybase.cma.ca/dbtw-wpd/Policypdf/PD18-01.pdf">http://policybase.cma.ca/dbtw-wpd/Policypdf/PD18-01.pdf</a> (accessed 12 Jan 2018).
- 82. Shanafelt T, Goh J, Sinsky C. The business case for investing in physician well-being. *JAMA Intern Med* 2017;177(2):195-205.
- 83. Gaba DM, Howard SK. Fatigue among clinicians and the safety of patients. *N Engl J Med* 2002;347(16):1249-55.
- 84. Ward S, Outram S. Medicine: in need of culture change. Itern Med J 2016;46(1):112-6.
- 85. Panagioti M, Panagopoulou E, Bower P, et al. Controlled interventions to reduce burnout in physicians: a systematic review and meta-analysis. *JAMA Intern Med* 2017;177(2):195-205.
- 86. American Medical Association. STEPSforward<sup>TM</sup>. Chicago, IL: American Medical Association, 2017. https://www.stepsforward.org/modules/joy-in-medicine (accessed 12 Jan 2018).
- 87. Jackson C. The Chalder Fatigue scale (CFQ 11). Occup Med 2015;65(1):86. doi: 10.1093/occmed/kqu168
- 88. Chalder T, Berelowitz G, Pawlikowska T, et al. Development of a fatigue scale. *J Psychosom Res* 1993;37(2):147-53.
- 89. Smets E, Garssen B, Bonke Bd, De Haes J. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *J Psychosom Res* 1995;39(3):315-25.
- 90. Ioannidis JPA, Greenland S, Hlatky MA, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;383(9912):166-75.
- 91. Gagliardi AR, Berta W, Kothari A, Boyko J, Urquhart R. Integrated knowledge translation (IKT) in health care: a scoping review. *Implement Sci* 2016;11:38. https://doi.org/10.1186/s13012-016-0399-1

BMJ Open: first published as 10.1136/bmjopen-2018-021967 on 21 September 2018. Downloaded from http://bmjopen.bmj.com/ on June 9, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

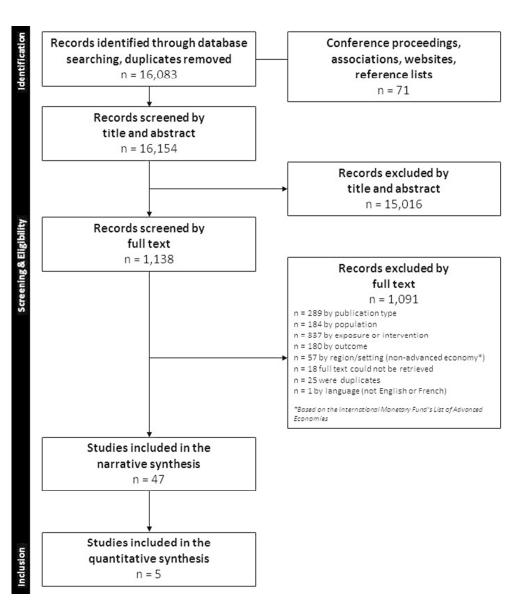


Figure 1. Flow of records through the selection process  $60\times69$ mm (300 x 300 DPI)

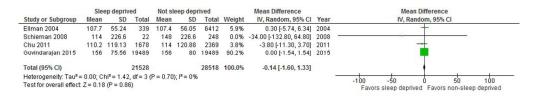
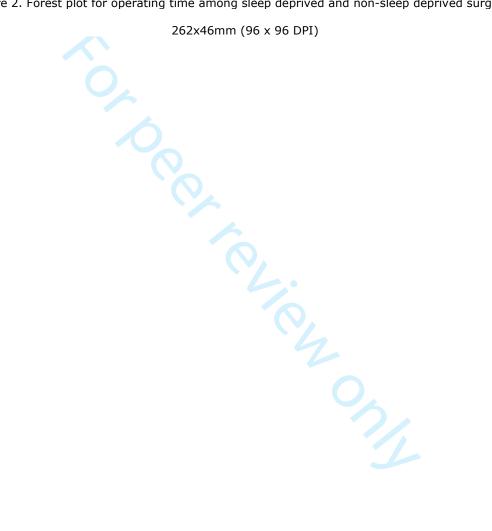


Figure 2. Forest plot for operating time among sleep deprived and non-sleep deprived surgeons



BMJ Open: first published as 10.1136/bmjopen-2018-021967 on 21 September 2018. Downloaded from http://bmjopen.bmj.com/ on June 9, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

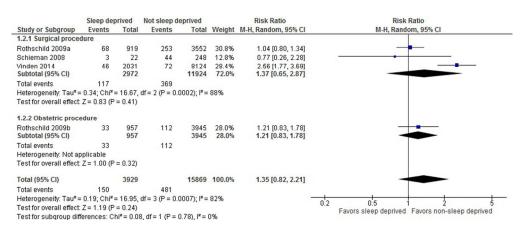
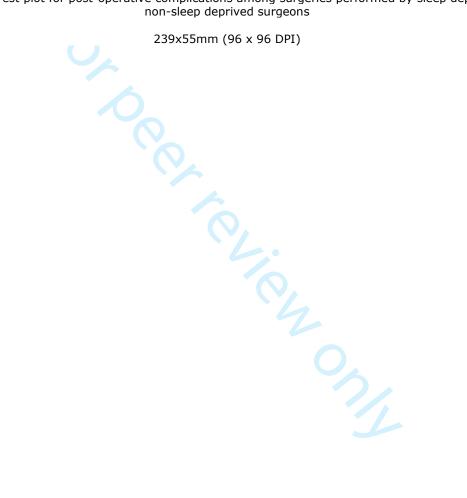


Figure 3. Forest plot for intra-operative complications among procedures performed by sleep deprived and non-sleep deprived surgeons or obstetrician-gynecologists

239x97mm (96 x 96 DPI)

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

Figure 4. Forest plot for post-operative complications among surgeries performed by sleep deprived and non-sleep deprived surgeons



BMJ Open: first published as 10.1136/bmjopen-2018-021967 on 21 September 2018. Downloaded from http://bmjopen.bmj.com/ on June 9, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES).

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

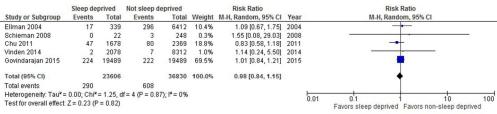
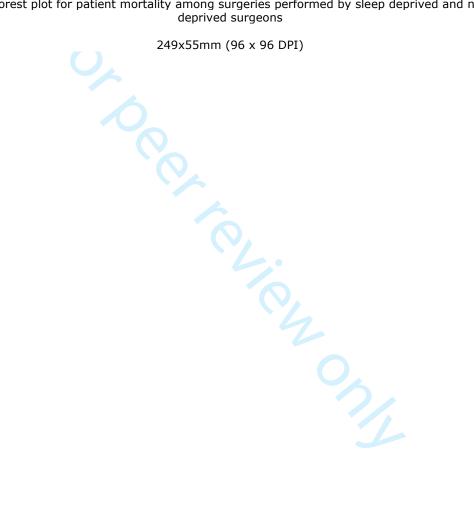


Figure 5. Forest plot for patient mortality among surgeries performed by sleep deprived and non-sleep deprived surgeons



Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

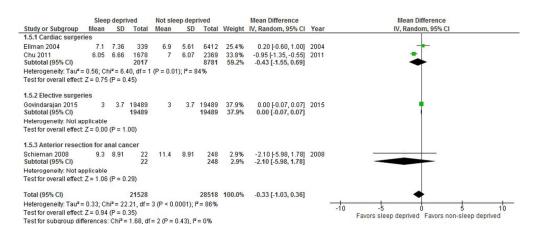
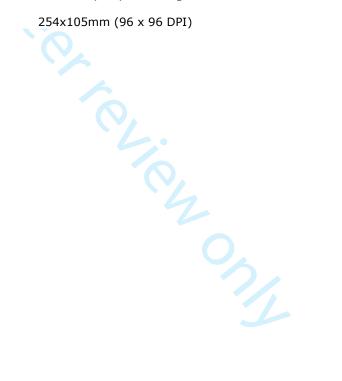


Figure 6. Forest plot for patient length of hospital stay (days) among surgeries performed by sleep deprived and non-sleep deprived surgeons



Database: In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1946 to Present

Date searched: 13 April 2016, updated 7 November 2017

Records retrieved: 5068 and 1442 in the update (removed duplicates retrieved in previous search)

- 1. Medical Staff, Hospital/
- 2. Physician Impairment/
- 3. exp Physicians/
- 4. allergist\*.ti.

- 5. (an?esthetist\* or an?esthesiologist\*).ti.
- 6. cardiologist\*.ti.
- 7. clinician\*.ti.
- 8. (clinician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 9. (clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw,kf.
- 10. dermatologist\*.ti.
- 11. endocrinologist\*.ti.
- 12. doctor\*.ti.
- 13. (doctor\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 14. (doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw,kf.
- 15. family practitioner\*.ti.
- 16. gastroenterologist\*.ti.
- 17. (general practitioner\* or GP\*).ti.
- 18. (general adj2 physician\*).ti.
- 19. geriatrician\*.ti.
- 20. gyn?ecologist\*.ti.

```
21. h?ematologist*.ti.
```

- 22. (health\* adj2 (professional\* or provider\*)).ti.
- 23. hospitalist\*.ti.
- 24. (house staff\* or housestaff\*).ti.
- 25. intensivist\*.ti.
- 26. internist\*.ti.
- 27. medical professional\*.ti.
- 28. obstetrician\*.ti.
- 29. oncologist\*.ti.
- 30. ophthalmologist\*.ti.
- 31. orthop?edist\*.ti.
- 32. (otolaryngologist\* or otorhinolaryngologist\*).ti.
- 33. neonatologist\*.ti.
- 34. nephrologist\*.ti.
- 35. neurologist\*.ti.
- 36. neuropsychiatrist\*.ti.
- 37. neurosurgeon\*.ti.
- 38. p?ediatrician\*.ti.
- 39. perinatologist\*.ti.
- 40. physician\*.ti.
- 41. (physician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 42. (physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw,kf.
- 43. primary care practitioner\*.ti.
- 44. psychiatrist\*.ti.
- 45. pulmonologist\*.ti.
- 46. rheumatologist\*.ti.
- 47. surgeon\*.ti.

- 48. (surgeon\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 49. traumatologist\*.ti.
- 50. urologist\*.ti.
- 51. or/1-50 [Combined MeSH, title, and text word searches for physicians]
- 52. Burnout, Professional/
- 53. exp Circadian Rhythm/
- 54. exp Fatigue/
- 55. Occupational Health/
- 56. Rest/ph, px [Physiology, Psychology]
- 57. Sleep Deprivation/
- 58. Sleep Disorders, Circadian Rhythm/
- 59. Sleep Wake Disorders/
- 60. exp Stress, Psychological/
- 61. Workload/px [Psychology]
- 62. Work Schedule Tolerance/
- 63. ((24 hour\* or 24 hr\* or twenty four hour\* or twentyfour hour\*) adj rhythm\*).tw,kf.
- 64. biological rhythm\*.tw,kf.
- 65. (burn out\* or burned out\* or burnt out\* or burnout\*).tw,kf.
- 66. circadian misalignment.tw,kf.
- 67. ((circadian or diurnam or ultradian) adj rhythm\*).tw,kf.
- 68. exhaust\*.tw,kf.
- 69. fatigu\*.tw,kf.
- 70. (sleep\* adj3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)).tw,kf.
- 71. tired\*.tw,kf.
- 72. weariness.tw,kf.
- 73. or/52-72 [Combined MeSH and text words for fatigue]
- 74. and/51,73 [Combined concepts for physicians and fatigue]
- 75. animals/ not (animals/ and humans/)
- 76. 74 not 75
- 77. (comment or editorial or letter).pt.

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

```
78. 76 not 77
```

- 79. limit 78 to yr="2000-Current"
- 80. limit 79 to (english or french)
- 81. remove duplicates from 80

Database: Ovid Embase 1996 to 2016 Week 15

Date searched: 13 April 2016

Records retrieved: 8859

- 1. medical staff/
- 2. exp physician/
- 3. allergist\*.ti.
- 4. (an?esthetist\* or an?esthesiologist\*).ti.
- 5. cardiologist\*.ti.
- 6. clinician\*.ti.
- 7. (clinician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 8. (clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 9. dermatologist\*.ti.
- 10. endocrinologist\*.ti.
- 11. doctor\*.ti.
- 12. (doctor\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 13. (doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 14. family practitioner\*.ti.
- 15. gastroenterologist\*.ti.
- 16. (general practitioner\* or GP\*).ti.
- 17. (general adj2 physician\*).ti.

- 19. gyn?ecologist\*.ti.
- 20. h?ematologist\*.ti.
- 21. (health\* adj2 (professional\* or provider\*)).ti.
- 22. hospitalist\*.ti.
- 23. (house staff\* or housestaff\*).ti.
- 24. intensivist\*.ti.
- 25. internist\*.ti.
- 26. medical professional\*.ti.
- 27. obstetrician\*.ti.
- 28. oncologist\*.ti.
- 29. ophthalmologist\*.ti.
- 30. orthop?edist\*.ti.
- 31. (otolaryngologist\* or otorhinolaryngologist\*).ti.
- 32. neonatologist\*.ti.
- 33. nephrologist\*.ti.
- 34. neurologist\*.ti.
- 35. neuropsychiatrist\*.ti.
- 36. neurosurgeon\*.ti.
- 37. p?ediatrician\*.ti.
- 38. perinatologist\*.ti.
- 39. physician\*.ti.
- 40. (physician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 41. (physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 42. primary care practitioner\*.ti.
- 43. psychiatrist\*.ti.
- 44. pulmonologist\*.ti.
- 45. rheumatologist\*.ti.
- 46. surgeon\*.ti.

- 47. (surgeon\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 48. traumatologist\*.ti.
- 49. urologist\*.ti.
- 50. or/1-49 [Combined Emtree, title, and text word searches for physicians]
- 51. burnout/
- 52. circadian rhythm/
- 53. circadian rhythm sleep disorder/
- 54. fatigue/
- 55. mental stress/
- 56. occupational health/
- 57. sleep deprivation/
- 58. sleep waking cycle/
- 59. work capacity/
- 60. work schedule/
- 61. working time/
- 62. workload/
- 63. ((24 hour\* or 24 hr\* or twenty four hour\* or twentyfour hour\*) adj rhythm\*).tw.
- 64. biological rhythm\*.tw.
- 65. (burn out\* or burned out\* or burnt out\* or burnout\*).tw.
- 66. circadian misalignment.tw.
- 67. ((circadian or diurnam or ultradian) adj rhythm\*).tw.
- 68. exhaust\*.tw.
- 69. fatigu\*.tw.
- 70. (sleep\* adj3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)).tw.
- 71. tired\*.tw.
- 72. weariness.tw.
- 73. or/51-72 [Combined Emtree and text words for fatigue]
- 74. and/50,73 [Combined concepts for physicians and fatigue]
- 75. animals/ not (animals/ and humans/)
- 76. 74 not 75

BMJ Open: first published as 10.1136/bmjopen-2018-021967 on 21 September 2018. Downloaded from http://bmjopen.bmj.com/ on June 9, 2025 at Agence Bibliographique de I

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

77. (conference\* or editorial or letter or proceeding).pt.

78. 76 not 77

79. limit 78 to yr="2000-Current"

80. limit 79 to (english or french)

81. limit 80 to embase

Database: Ovid PsycINFO 1987 to April Week 1 2016

Date searched: 13 April 2016

Records retrieved: 2094

- 1. exp Physicians/
- 2. allergist\*.ti.
- 3. (an?esthetist\* or an?esthesiologist\*).ti.
- 4. cardiologist\*.ti.
- 5. clinician\*.ti.
- 6. (clinician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 7. (clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 8. dermatologist\*.ti.
- 9. endocrinologist\*.ti.
- 10. doctor\*.ti.
- 11. (doctor\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 12. (doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 13. family practitioner\*.ti.
- 14. gastroenterologist\*.ti.
- 15. (general practitioner\* or GP\*).ti.
- 16. (general adj2 physician\*).ti.

```
17. geriatrician*.ti.
```

- 18. gyn?ecologist\*.ti.
- 19. h?ematologist\*.ti.
- 20. (health\* adj2 (professional\* or provider\*)).ti.
- 21. hospitalist\*.ti.
- 22. intensivist\*.ti.
- 23. internist\*.ti.
- 24. medical professional\*.ti.
- 25. obstetrician\*.ti.
- 26. oncologist\*.ti.
- 27. ophthalmologist\*.ti.
- 28. orthop?edist\*.ti.
- 29. (otolaryngologist\* or otorhinolaryngologist\*).ti.
- 30. neonatologist\*.ti.
- 31. nephrologist\*.ti.
- 32. neurologist\*.ti.
- 33. neuropsychiatrist\*.ti.
- 34. neurosurgeon\*.ti.
- 35. p?ediatrician\*.ti.
- 36. perinatologist\*.ti.
- 37. physician\*.ti.
- 38. (physician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 39. (physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 40. primary care practitioner\*.ti.
- 41. psychiatrist\*.ti.
- 42. pulmonologist\*.ti.
- 43. rheumatologist\*.ti.
- 44. surgeon\*.ti.

- 46. traumatologist\*.ti.
- 47. urologist\*.ti.

- 48. or/1-47 [Combined thesaurus, title, and text word searches for physicians]
- 49. Compassion Fatigue/
- 50. Fatigue/
- 51. Human Biological Rhythms/
- 52. Occupational Health/
- 53. Occupational Stress/
- 54. Sleep/
- 55. Sleepiness/
- 56. Working Conditions/
- 57. Work Rest Cycles/
- 58. Work Week Length/
- 59. Work Scheduling/
- 60. Workday Shifts/
- 61. ((24 hour\* or 24 hr\* or twenty four hour\* or twentyfour hour\*) adj rhythm\*).tw.
- 62. biological rhythm\*.tw.
- 63. (burn out\* or burned out\* or burnt out\* or burnout\*).tw.
- 64. circadian misalignment.tw.
- 65. ((circadian or diurnam or ultradian) adj rhythm\*).tw.
- 66. exhaust\*.tw.
- 67. fatigu\*.tw.
- 68. (sleep\* adj3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)).tw.
- 69. tired\*.tw.
- 70. weariness.tw.
- 71. or/49-70 [Combined thesaurus and text words for fatigue]
- 72. and/48,71 [Combined concepts for physicians and fatigue]
- 73. limit 72 to yr="2000-Current"
- 74. limit 73 to (english or french)

Database: CINAHL Plus with Full Text (1937 to the present) via EBSCOhost

Date searched: 14 April 2016

Records retrieved: 3378

- S1. (MH "Medical Staff, Hospital+")
- S2. (MH "Physicians+")
- S3. TI allertist\*
- S4. TI (anesthetist\* or anaesthetist\* or anesthesiologist\* or anaesthesiologist\*)
- S5. TI cardiologist\*
- S6. TI clinician\*
- S7. clinician\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or "work\* hour\*" or "work life balance")
- S8. clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)
- S9. TI dermatologist\*
- S10. TI endocrinologist\*
- S11. TI doctor\*
- S12. doctor\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or work\* hour\* or "work life balance")
- S13. doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)
- S14. TI "family practitioner\*"
- S15. TI gastroenterologist\*
- S16. TI ("general practitioner\*" or GP\*)
- S17. TI (general N2 physician\*)
- S18. TI geriatrician\*
- S19. TI (gynaecologist\* or gynecologist\*)
- S20. TI (haematologist\* or hematologist\*)
- S21. TI hospitalist\*
- S22. TI ("house staff\*" or housestaff\*)

S24. TI internist\*

- S25. TI obstetrician\*
- S26. TI oncologist\*
- S27. TI ophthalmologist\*
- S28. TI (orthopaedist\* or orthopedist\*)
- S29. TI (otolaryngologist\* or otorhinolaryngologist\*)
- S30. TI neonatologist\*
- S31. TI nephrologist\*
- S32. TI neurologist\*
- S33. TI neuropsychiatrist\*
- S34. TI neurosurgeon\*
- S35. TI (paediatrician\* OR pediatrician\*)
- S36. TI perinatologist\*
- S37. TI physician\*
- S38. physician\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or "work\* hour\*" or "work life balance")
- S39. physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)
- S40. TI "primary care practitioner\*"
- S41. TI psychiatrist\*
- S42. TI pulmonologist\*
- S43. TI rheumatologist\*
- S44. TI surgeon\*
- S45. surgeon\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or work\* hour\* or "work life balance")
- S46. TI traumatologist\*
- S47. TI urologist\*
- S48. S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41 OR

S42 OR S43 OR S44 OR S45 OR S46 OR S47

- S49. (MH "Circadian Rhythm")
- S50. (MH "Fatigue")
- S51. (MH "Impairment, Health Professional")
- S52. (MH "Mental Fatigue")
- S53. (MH "Occupational Health")
- S54. (MH "Shiftwork")
- S55. (MH "Sleep Deprivation")
- S56. (MH "Sleep Disorders, Circadian Rhythm")
- S57. (MH "Sleep-Wake Transition Disorders")
- S58. (MH "Stress, Occupational+")
- S59. (MH "Stress, Psychological")
- S60. ("24 hour\*" or "24 hr\*" or "twenty four hour\*" or "twentyfour hour\*") N1 rhythm\*
- S61. "biological rhythm\*"
- S62. "burn out\*" or "burned out\*" or "burnt out\*" or burnout\*
- S63. "circadian misalignment"
- S64. (circadian or diurnam or ultradian) N1 rhythm\*
- S65. exhaust\*
- S66. fatigu\*
- S67. sleep\* N3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)
- S68. tired\*
- S69. weariness
- S70. S49 OR S50 OR S51 OR S52 OR S53 OR S54 OR S55 OR S56 OR S57 OR S58 OR S59 OR S60 OR S61 OR
- S62 OR S63 OR S64 OR S65 OR S66 OR S67 OR S68 OR S69
- S71. S48 AND S70
- S72. S48 AND S70 Limiters Published Date: 20000101-20161231; Publication Type: Clinical Trial, Journal

Article, Meta Analysis, Meta Synthesis, Practice Guidelines, Randomized Controlled Trial, Research,

Review, Systematic Review; Language: English, French

"family practitioners"[ti] OR gastroenterologist[ti] OR gastroenterologists[ti] OR "general practice physician"[ti] OR "general practice physicians"[ti] OR "general practitioner"[ti] OR "general practitioners"[ti] OR geriatrician[ti] OR geriatricians[ti] OR gynaecologist[ti] OR gynaecologists[ti] OR gynecologist[ti] OR gynecologists[ti] OR haematologist[ti] OR haematologists[ti] OR hematologist[ti] OR hematologists[ti] OR "health care professional"[ti] OR "health care professionals"[ti] AND "health care provider"[ti] OR "health care providers" OR "health professional"[ti] OR "health professionals"[ti] OR "health provider"[ti] OR "health providers"[ti] OR "healthcare professional"[ti] OR "healthcare professionals"[ti] OR "healthcare provider"[ti] OR "healthcare providers"[ti] OR hospitalist[ti] OR hospitalists[ti] OR "house staff"[ti] OR "house staffs"[ti] OR housestaff[ti] OR housestaffs[ti] OR intensivist[ti] OR intensivists[ti] OR internist[ti] OR internists[ti] OR "medical professional"[ti] OR "medical professionals"[ti] OR obstetrician[ti] OR obstetricians[ti] OR oncologist[ti] OR oncologists[ti] OR ophthalmologist[ti] OR ophthalmologists[ti] OR orthopaedist[ti] OR orthopaedists[ti] OR orthopaedist[ti] OR orthopedists[ti] OR otolaryngologist[ti] OR otolaryngologists[ti] OR otorhinolaryngologist[ti] OR otorhinolaryngologists[ti] OR neonatologist[ti] OR neonatologists[ti] OR nephrologist[ti] OR nephrologists[ti] OR neurologist[ti] OR neurologists[ti] OR neuropsychiatrist[ti] OR neuropsychiatrists[ti] OR neurosurgeon[ti] OR neurosurgeons[ti] OR paediatrician[ti] OR paediatricians[ti] OR pediatrician[ti] OR pediatricians[ti] OR perinatologist[ti] OR perinatologists[ti] OR physicians[ti] OR physicians[ti] OR ((physician[tiab] OR physicians[tiab]) AND (absent[tiab] OR absentee[tiab] OR absenteeism[tiab] OR absenteeisms[tiab] OR "burned out"[tiab] OR "burn out"[tiab] OR "burn outs"[tiab] OR burnout[tiab] OR burnouts[tiab] OR "burnt out"[tiab] OR coping[tiab] OR distress[tiab] OR distressed[tiab] OR distresses[tiab] OR distressing[tiab] OR "duty hour"[tiab] OR "duty hours"[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR health[tiab] OR healthy[tiab] OR impair[tiab] OR impaired[tiab] OR impairment[tiab] OR resilience[tiab] OR resiliency[tiab] OR resilient[tiab] OR satisfaction[tiab] OR sleep[tiab] OR sleepiness[tiab] OR sleeps[tiab] OR "well being"[tiab] OR wellbeing[tiab] OR wellness[tiab] OR "work hours"[tiab] OR "working hours"[tiab] OR "work life balance"[tiab])) OR ((physician[tiab] OR physicians[tiab]) AND (cities[tiab] OR city[tiab] OR community[tiab] OR communities[tiab] OR country[tiab] OR countryside[tiab] OR frontier[tiab] OR north[tiab] OR northern[tiab] OR remote[tiab] OR rural[tiab] OR suburb[tiab] OR suburbs[tiab] OR suburban[tiab] OR urban[tiab] OR urbanite[tiab])) OR "primary care practitioner"[ti] OR "primary care practitioners"[ti] OR psychiatrist[ti] OR psychiatrists[ti] OR pulmonologist[ti] OR pulmonologists[ti] OR rheumatologist[ti] OR rheumatologists[ti] OR surgeon[ti] OR surgeons[ti] OR ((surgeon[tiab] OR surgeons[tiab]) AND (absent[tiab] OR absentee[tiab] OR absenteeism[tiab] OR absenteeisms[tiab] OR

"burned out"[tiab] OR "burn out"[tiab] OR "burn outs"[tiab] OR burnout[tiab] OR burnouts[tiab] OR "burnt out"[tiab] OR coping[tiab] OR distress[tiab] OR distressed[tiab] OR distresses[tiab] OR distressing[tiab] OR "duty hour"[tiab] OR "duty hours"[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR health[tiab] OR healthv[tiab] OR impair[tiab] OR impaired[tiab] OR impairment[tiab] OR resilience[tiab] OR resiliency[tiab] OR resilient[tiab] OR satisfaction[tiab] OR sleep[tiab] OR sleepiness[tiab] OR sleeps[tiab] OR "well being"[tiab] OR wellbeing[tiab] OR wellness[tiab] OR "work hours"[tiab] OR "working hours"[tiab] OR "work life balance"[tiab])) OR traumatologist[ti] OR traumatologists[ti] OR urologist[ti] OR urologists[ti]) AND ("Burnout, Professional"[mh:noexp] OR "Circadian Rhythm"[mh] OR "Fatigue"[mh] OR "Occupational Health"[mh:noexp] OR "Rest/physiology"[mh:noexp] OR "Rest/psychology"[mh:noexp] OR "Sleep Deprivation"[mh:noexp] OR "Sleep Disorders, Circadian Rhythm"[mh:noexp] OR "Stress, Psychological"[mh] OR "Workload/psychology"[mh] OR "Work Schedule Tolerance"[mh:noexp] OR "24 hour rhythm"[tiab] OR "24 hour rhythms"[tiab] OR "24 hr rhythm"[tiab] OR "24 hr rhythms"[tiab] OR alertness[tiab] OR "biological rhythm"[tiab] OR "biological rhythms"[tiab] OR "burn out"[tiab] OR "burned out"[tiab] OR "burnt out"[tiab] OR burnout[tiab] OR "circadian misalignment"[tiab] OR "circadian rhythm"[tiab] OR "circadian rhythms"[tiab] OR "diurnal rhythm"[tiab] OR "diurnal rhythms"[tiab] OR exhausted[tiab] OR exhaustion[tiab] OR exhausting[tiab] OR exhausts[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR (("Sleep"[mh:noexp] OR sleep[tiab] OR sleeping[tiab]) AND (deprivation[tiab] OR deprive[tiab] OR deprived[tiab] OR deprives[tiab] OR depriving[tiab] OR disorder[tiab] OR disorders[tiab] OR lack[tiab] OR lacked[tiab] OR lacking[tiab] OR lacks[tiab] OR loss[tiab] AND insufficient[tiab] OR problem[tiab] OR problems[tiab])) OR tired[tiab] OR tiredness[tiab] OR "twenty four hour rhythm"[tiab] OR "twenty four hour rhythms"[tiab] OR weariness[tiab] OR "ultradian rhythm"[tiab] OR "ultradian rhythms"[tiab])) NOT (((Animals[MESH] OR Animal Experimentation[MESH] OR "Models, Animal"[MESH] OR Vertebrates[MESH]) NOT (Humans[MESH] OR Human experimentation[MESH])) OR (((animals[tiab] OR animal model[tiab] OR rat[tiab] OR rats[tiab] OR mouse[tiab] OR mice[tiab] OR rabbit[tiab] OR rabbits[tiab] OR pig[tiab] OR pigs[tiab] OR porcine[tiab] OR swine[tiab] OR dog[tiab] OR dogs[tiab] OR hamster[tiab] OR hamsters[tiab] OR chicken[tiab] OR chickens[tiab] OR sheep[tiab]) AND (publisher[sb] OR inprocess[sb] OR pubmednotmedline[sb])) NOT (human[ti] OR humans[ti] OR people[ti] OR children[ti] OR adults[ti] OR seniors[ti] OR patient[ti] OR patients[ti])))) NOT (editorial[pt] OR comment[pt] OR letter[pt] OR newspaper article[pt])) AND ((publisher[sb] NOT pubstatusnihms NOT pubstatuspmcsd NOT pmcbook) OR (pubstatUSheadofprint))

## **Supplementary table 1.** Descriptive characteristics of the included studies

				BMJ O <sub>f</sub>	5/bmjopen-2018-021967 coted by copyright, inducention			
Supplementary	table 1. Descriptive char	acterist	ics of the inc	luded studies			8-021: ight, i	
Study	Physician and patient chara	cteristics			Setting		Interventions or 5	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban	exposures ding 2	
						or	<b>→</b> —	
						rural	for l	
Observational (exp	oosure) studies (n=45)						pte En	
Cohort design							mb sei s r	
Chu, 2011 [32]	Surgeons	6	NR	Range: 32-55y	Tertiary care academic	Urban	Sleep restriction Sue	Length of surgery; patient
Canada	Patients: cardiac surgery	4,047	NR	NR	hospital		to work on 🛱 👼 🥦	postoperative mortality,
	cases						preceding sægger	complications, length of
							t Su tex	stay
Ellman, 2004 [41]	Surgeons	NR	NR	NR	University hospitals	Urban	Sleep restrice de La Company	Length of surgery; patient
US	Patients: adult cardiac	6,751	70%	S: 63.4±0.7y	<del>-</del>		to work on the he	complications, in-hospital
	surgery cases			C: 63.5±0.1y			preceding support	mortality, length of stay,
							a m	need for blood products
Govindarajan,	Surgeons	1,448	NR	46.3±8.7	Academic and non-	Mixed	Sleep deprivation	Length of surgery; Patient
2015 [31]	Patients: surgical cases	38,978	NR	56.4±16.6y	academic hospitals		to work on 🗖 e n 🔁 ht	complications, mortality,
Canada							preceding a <b>≧</b> ayt <mark>m</mark> e	readmissions, length of
							surgery 💆 🧵	stay
Rothschild, 2009	Surgeons	220	Surgeons:	Surgeons:	Tertiary care academic	Urban	Sleep deprivetion due	Patient complications,
[58]	Obstetrician/gynecologists		84%	42.0±7.6y	trauma centre/referral		to work on <b>G</b> e n <b>g</b> ht	preventable
US			OB/GYNs:	OB/GYNs:	centre for high-risk		preceding a aytime	complications
			28%	42.0±9.0y	obstetrics		procedure procedure	
	Patients: surgical and	Surg.:	Surg:	Surg:			com/ on June 9, similar technolo	
	obstetrics cases	4,471	S: 25%	S: 49.1±16.3y			ar t	
		Obst.:	C: 28%	C: 50.0±16.3y			Jul	
		4,902	Obst.:	Obst.:			ne (	
			S: 0%	S: 32.9±5.2y			9, 20 logie	
			C: 0%	C: 33.5±5.0y			Fatigue due to wark	
Schieman, 2007	Colorectal surgeons	NR	NR	NR	University teaching	NR	<u> </u>	Length of surgery; patient
[63]					hospitals		Ó	operative complications,
Canada	Patients: undergoing	270	NR	S: 64.5y			surgery	length of stay, mortality,
	anterior resection for rectal			C: 64.4y			e <del>-</del>	cancer recurrence
	cancer						3ib	
							surgery ence Bibliographique de	
							òhic	<u>-</u>
							ļue	<u>-</u>
	For	peer rev	iew only - htt	p://bmjopen.b	mj.com/site/about/gu	idelines	xhtml 🔓	

				6/bmjopen-2018- cted by copyrig				
Study	Physician and patient chara	cteristics			Setting		Interventions or 2	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	including exposures	
Vinden, 2014 [77] Canada	General surgeons Patients: Elective cholecystectomies	331 10,390	83% S: 27% C: 26%	48±10y S: 49±16y C: 49±16y	Community hospitals	Mixed	Sleep deprivation value to overnight with the preceding deprivation of the surgery of the surger	Patient mortality, operative complications
<b>Before-after desigr</b> Amirian, 2014 [34] Denmark		29	55%	Median: 35y Range: 27-49y	Academic hospital	Urban	17-h night shift with sleep deprivation to the	Cognitive and psychomotor abilities on a laparoscopic simulation
Gerdes, 2008 [45] US	Surgeons	9	NR	NR	University Hospital	Urban	Fatigue; sleep in a deprivation a fatigue overnight can shift	Cognitive and psychomotor abilities
Lederer, 2006 [50] Austria	Senior anesthetists	11	82%	49.0±2.0y	Hospital	Urban	Sleep deprive to shape from 24-h can be shaped by the shape from 24-h can be shaped by the shaped by	Concentration ability; reaction time; performance on psychometric tasks
Time series design							ain Jo	
Leichtfried, 2011 [51] Austria	Anesthetists	10	100%	Mean: 32y Range: 29-35y	University Hospital	Urban	Sleep deprivations from 24-h shift; sleepiness, Reepiness	Melatonin metabolite profile
Cross-sectional des	sign						nilar o	
Aziz, 2004 [35] US	Family medicine physicians Various specialties	153	NR	NR	Hospitals	NR	Fatigue # 3	Stress
Beaujouan, 2005 [36] France	Anesthesiologists	3,476	64%	≤35y: 9% 36-45y: 28% 46-55y: 49% 56-65y: 13%	Public sector General hospitals University hospitals Private hospitals	NR	June 9, 2025 at Age	Substance abuse
Chang, 2013 [37] US	Anesthesiologists	11	64%	Mean: 38y IQR: 34-48y	Level 1 trauma centre	NR	Sleep deprivation due	Cognitive performance; reaction time
	For	peer rev	iew only - htt	p://bmjopen.b	mj.com/site/about/gu	uidelines	shift; sleepiness bill ographique de .xhtml	

			cted by copyrighting interventions					
Study	Physician and patient chara	cteristics			Setting	rights or 21	Outcomes	
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	967 on 2 including exposures	
Chen, 2008 [38] US	Psychiatrists Internists General practitioners Surgeons Obstetrician-gynecologists Radiologists Pediatricians Other	180	77%	Academic: 79% 36-55y Private practice: 73% 36-65y	Medical school Private practices	Urban	Sleep deprive uses related to text a	Impact on personal and professional life; perceived risk of errors
Doppia, 2011 [39] France	Anesthesiologists	565	64%	<35y: 11% 35-54y: 63% >55y: 25%	Public hospitals Private hospitals Work-health environments Public health units	NR	Sleep deprive data minin	Burnout
Elovaino, 2015 [42] Finland	Physicians in various specialties	1,524	40%	Median: 49.7y Range: 24-69y	Hospitals Primary care Private practice Other unspecified	NR	Sleep difficult training	Job demands and contro
Gander, 2000 [43] New Zealand	Anesthetists	183	NR	Mean: 46y	Combined public/private practice Other unspecified	NR	Work-related sleep disturbanced Si	Risk of fatigue-related errors
Harbeck, 2015 [46] Germany	Internists	20	45%	Median: 32y Range: 26-42y	Hospital	NR	Sleep disturtiance due to a 24-call shift on Jun	Biochemical and physiological parameters neurocognitive function
Heponiemi, 2014 [47] Finland	Physicians in various specialties Non-specialized physicians	1,541	40%	49.80±9.49y, Range: 24-67y	Hospitals Primary care clinic Private practice Other unspecified	NR	Sleep difficuties 🖁	Job satisfaction; work ability; psychological distress
	For	neer rev	iew only - htt:	o://hmionen.h	mj.com/site/about/gu	idelines	2025 at Agence Bibliographique de I gies.	

				ВМЈ О	oen		5/bmjopen-2018-0219 cted by copyrightsin	
Study	Physician and patient chara	cteristics			Setting		Interventions or 2	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures cluding 67 on 2.	
Jackson, 2017 [48] US	Surgeons in various subspecialties	993	61%	More; less satisfied: 30-39y: 23%;24% 40-49y: 32%;36% 50-59y: 23%;27% ≥60y: 23%;14%	Academic practice Non-academic practice	NR	Beeptember 2018. Downloader feeling uses related to text and da	
Kanieta, 2011 [49] Japan	Internists Surgeons Orthopedics Pediatricians Obstetrician-gynecologists Psychiatrists Dermatologists Urologists Opthalmologists Otorhinolaryngologists Other	3,486	66%	20-39y: 11% 40-49y: 25% 50-59y: 28% 60-69y: 16% ≥70y: 21%	Hospitals Clinics Other unspecified	NR	Sleep deprive mile difficulties; Manual Simile difficulties; Al training, and simile difficulties; Al training, and simile difficulties.	Medical incidents
Lindfors, 2006 [52] Finland	Anesthetists	328	53%	47±7.8y Range: 32-69y	University hospitals Central and district hospitals Private sector	NR	Sleep disturbances; sleepiness nolog	Stress; suicidal tendencie
Mahmood, 2016 [53] Norway	Generalists Internists Pediatricians Surgical specialties Anesthesiologists	450 (all time points)	41%	43y±2.8y	Public health system Private practice	NR	Sleep deprivations at Agence Bibliographique de l	Alcohol misuse
	Eor	neer revi	iew only - b++	n·//hmianen h	mj.com/site/about/gu	idalinas	raphique d	

			5/bmjopen-2018- cted by copyrig					
Study	Physician and patient cl	haracteristics			Setting		right -021	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	- 5 0	
Nishimura, 2014 54] apan	Neurosurgeons and neurologists	2,564	NR	NR	Stroke care centres Teaching hospitals	NR	Sleep deprivent Ense	Burnout
Pit, 2014 [55] Australia	General practitioners	92	60%	50±10.7y	NR	Rural	Work-related disturbance of 20	Early retirement intentions
Pit, 2016 [56] Australia	General practitioners	92	60%	50±10.7y	Private (solo) practice Group practice	Rural	Work-related \$\frac{4}{2} \frac{4}{2} \fra	Sickness presenteeism
Roberts, 2014 [57] US	General internists Internal medicine hospitalists	578	58%	Hospitalists: 46.9±12.4y Generalists: 53.6±10.2y	Private practice Academic medical centre Veterans hospital Military practice Other	NR	wnloaded from ht uperieur (ABES) . xt and data mining	Falling asleep while driving
Gaadat, 2016 [60] JS	Anesthesiologists	21	71%	30-40y: 57% 41-50y: 19% 51-55y: 24% Range: 32-56y	Tertiary care academic children's hospital	NR	Sleep deprivation due to 17-h night call hift ain open g, a	Mood disturbances
Saadat, 2017 [59] JS	Anesthesiologists	21	65%	Range: 32-56 years	Tertiary care academic children's hospital	NR	Sleep deprivation due to 17-h nigh call hift	Reaction time
anches, 2015 [61] pain	Emergency medicine physicians	18	28%	29.2±2.6y	Central hospital	NR	Sleep deprivation	Cognitive and psychomotor abilities
argent, 2009 [62] JS	Orthopedic surgeons	264	92%	NR	Orthopedic surgery training programs	NR	Sleep deprivatione 9, 20	Burnout; psychological distress; marital satisfaction
ende, 2012 [64] rance	Emergency physicians	318	62%	39±8y	Hospitals Mobile emergency services Other unspecified	NR	Fatigue; sle	Stress
		For peer revi	ew only - httl	o://bmjopen.b	mj.com/site/about/gu	idelines	deprivation at Agence Bibliographique de I	

		cted by copyright or	bm jopen-2018						
Study	Physician and patient cha	aracteristics			Setting	<u>ရှိ</u> Intervention <u>s</u> or	8-021	Outcomes	
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures C C C Si	967 on 21	
Sexton, 2001 [65] US	Consulting physicians: Surgeons Anesthesiologists Pulmonary physicians Cardiologists Pediatricians	271	NR	NR	Teaching and non- teaching hospitals	Urban		1 September 2018.	Perceived performance effectiveness
Shanafelt, 2005 [67] US, Canada, Mexico	Oncologists	241	85%	>50y: 51%	Community clinics Hospitals Private practice Academic medical centres	NR	Fatigue; sleeps t Superieur (	Downloaded	Quality of life/well-being
Shanafelt, 2010 [66] US	Surgeons	7,905	87%	Median: 51y Q1: 43y Q2: 59y	Private practice Academic medical centres Veterans hospital Active military practice Retired or not in practice Other	NR	Fatigue Fatigue Fatigue Fatigue Fatigue	rom http://bmjopen.b	Perceived major medical errors
Shanafelt, 2014 [68] US	Oncologists	1,117	52%	Median: 52y	Private practice Academic practice Veteran's hospital Industry, other	NR		mi.com/ on	Satisfaction with work-life balance
Shirom, 2006 [69] Israel	Opthalmologists Dermatologists Otolaryngologists Gynecologists General surgeons Cardiologists	890	80%	Median: 52y SD: 7.2y	Community clinics Acute care hospital outpatient clinics	NR	Physical fatishe	June 9, 2025 at Agence Bibliographique de l	Perception of quality of patient care
	F	or peer rev	iew only - htt	p://bmjopen.k	omj.com/site/about/gui	idelines.	xhtml	draphique de l	

			5/bmjopen-2018-0219 cted by copyrightisin					
Study	Physician and patient char	acteristics			Setting		Interventions or 21	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures exposures	
Shirom, 2010 [70] Israel	Opthalmologists Dermatologists Otolaryngologists Gynecologists General surgeons Cardiologists	890	80%	Median: 52y SD: 7.2y	Community clinics Acute care hospital outpatient clinics	NR	Physical fatiar uses related to	Burnout
Smith, 2017 [71] UK	General practitioners Surgeons Other unspecified specialties	3,550	63%	NR	NR (varied)	NR	Perceived fattigues sleep deprivation de	Physical and mental health; competence
Starmer, 2016 [72] US	General pediatricians Pediatric surgeons Pediatric hospitalists Pediatric specialists (unspecified)	840	40%	NR	NR (some in private practice)	NR	Sleep deprivemining, Al training, and Fatigue	Burnout; balanced personal and professional commitments; life and career satisfaction
Tanti, 2017 [73] Malta	Physicians (unspecified)	204	62%	Median: 41y	Hospitals Community Office-based	NR	Fatigue aining, a	Prescribing errors
Tokuda, 2009 [74] Japan	Hospital physicians: Generalists Other unspecified specialties	236	75%	40.9±7.8y Range: 26-76y	Hospitals with ≥20 inpatient beds	NR	Sleep deprivation/on	Burnout; job satisfaction
Vela-Bueno, 2008 [76] Spain	Primary care physicians	113	27%	41.4±8.0y	Primary care centres	Urban	Sleep problems, ne 9, 20	Burnout
Wada, 2010 [78] Japan	Physicians (unspecified)	3,862	78%	M: 75% 30- 59y F: 85% 30-59y	Hospitals	NR	025 at Agence Bibliographic	Depressive symptoms
							ographi	

				ВМЈО	pen		6/bmjopen-2018 cted by copyrig	
Study	Physician and patient cha	ractoristics			Setting		면 역 8-0 Interventions or 21	Outcomes
Country	Type	n=	Sex (% male)	Δσρ	Location	Urban	- 5 0	Outcomes
country	Type	.,-	Sex (70 maie)	780	Location	or		
						rural	on 2	
Non-comparative of	design						<u> </u>	
Sander, 2008 [43]		20	85%	Median: 44y	Hospitals	Urban	Sleep distur ang	Psychomotor
New Zealand				,	,		from consecutive	performance
							working day 2 cc sqn-	<b>,</b>
							call work te	
ntervention studie	es (n=2)						<u>~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ </u>	
andomized contro							o te	
	Emergency physicians	17	35%	39.1y±6.9y	University hospital	Urban		Perceived stress; urine
rance	G -/ [- /			, ,	-1 6		and 24-h sh	interleukine-8
							deprivation a deep	
							quality; $\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{j=1}^{n} \sum_{j=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{j=1}^{n} \sum_{j=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{j=1}^$	
Jchal, 2005 [75]	Surgeons	64	67%	Median:	Government hospitals	NR	Sleep deprive	Product quality,
lorway	Gynecologists	O-I	0770	Post-call:	Government nospituis	1411	to 24-h call mit	procedure effectiveness
·O. way	Orthopedic surgeons			33.0y			7 0	of a surgical simulation
	Urologists			Post-work:			A ∯	or a surfical simulation
	Vascular surgeons			38.0y			bmjo train	
ilitea Kingaom; Os	5: United States of America,	y. year(s)					bmj.com/ on June 9, 2025 at Agendand similar technologies.	
	F	or peer rev	iew only - htt	p://bmjopen.b	omj.com/site/about/gu	idelines.	2025 at Agence Bibliographique de l gies.	

## **Supplementary file 3.** Risk of bias assessments

Summary of risk of bias assessments for randomized controlled trials (n=2)<sup>a</sup>

First Author, Year	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other sources of bias	Overall risk of bias <sup>b</sup>
Dutheil, 2013	Low	Unclear	High	High	Low	Low	High	High
Uchal, 2005	Low	Low	Unclear	Low	Low	Low	Low	Unclear

<sup>&</sup>lt;sup>a</sup>Assessed using the Cochrane Collaboration's Risk of Bias Tool

Summary of quality assessments for cohort studies (n=6)<sup>a</sup>

First Author,	Selection					Comparab	oility		Outco	me		Total
Year	Representa- tiveness of exposed cohort /1	Selection of non- exposed cohort /1	Ascertain- ment of exposure /1	Outcome not present at start /1	Total /4	Compara -bility of cohorts /2	Total /2	Assess- ment of outcome /1	Adequate length of follow-up /1	Adequate follow-up of cohorts /1	Total /1	Score <sup>b</sup> /9
Chu, 2011	1	1	0	1	3	2	2	1	1	1	3	8
Ellman, 2004	1	1	1	1	4	1	1	1	1	1	3	8
Govindarajan, 2015	1	1	1	1	4	2	2	1	1	1	3	9
Rothschild, 2009	1	1	1	1	4	2	2	1	1	1	3	9
Schieman, 2008	1	1	1	1	4	1	1	1	1	1	3	8
Vinden, 2014	1	1	1	1	4	1	1	1	1	1	3	8

<sup>&</sup>lt;sup>a</sup>Assessed using the Newcastle-Ottawa Quality Assessment Scale

<sup>&</sup>lt;sup>b</sup>Overall risk of bias is Low if all domains are rated as low, High if at least one domain is assessed as high, and Unclear if at least one domain is assessed as unclear and no domains are assessed as high

<sup>&</sup>lt;sup>b</sup>An overall score of 7 to 9 stars is considered as low risk of bias, 4 to 6 as unclear risk of bias, and 3 or less as high risk of bias

Summary of risk of bias assessments for before-after studies (n=3)<sup>a</sup>

First Author, Year	Random sequence generation <sup>b</sup>	Allocation concealment <sup>b</sup>	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other sources of bias <sup>c</sup>	Overall risk of bias <sup>d</sup>
Amirian, 2014	NA	NA	High	High	Low	Low	High	High
Gerdes, 2008	NA	NA	High	High	Low	Low	High	High
Lederer, 2006	NA	NA	High	High	Low	Low	High	High

<sup>&</sup>lt;sup>a</sup>Assessed using Cochrane Effective Practice and Organization of Care (EPOC) Review Group's criteria for before-after studies, adapted from the Cochrane Collaboration Risk of Bias Tool

Summary of risk of bias assessments for time series studies (n=1)<sup>a</sup>

First Author, Year	Intervention independent of other changes	Intervention effect pre- specified	Intervention unlikely to affect data collection	Allocation concealment <sup>a</sup>	Incomplete outcome data	Selective reporting	Other sources of bias <sup>c</sup>	Overall risk of bias <sup>d</sup>
Leitchfried, 2011	Low	High	Low	NA	Low	Low	High	High

<sup>&</sup>lt;sup>a</sup>Assessed using Cochrane Effective Practice and Organization of Care (EPOC) Review Group's criteria for interrupted time series studies, adapted from the Cochrane Collaboration Risk of Bias Tool

<sup>&</sup>lt;sup>b</sup>Assessed as 'not applicable' (NA) when the studies did not include a control group

<sup>&</sup>lt;sup>c</sup>Assessed as High due to lack of a control group

<sup>&</sup>lt;sup>d</sup>Overall risk of bias is Low if all domains are rated as low, High if at least one domain is assessed as high, and Unclear if at least one domain is assessed as unclear and no domains are assessed as high

<sup>&</sup>lt;sup>b</sup>Assessed as not applicable (NA) when the studies did not include a control group

<sup>&</sup>lt;sup>c</sup>Assessed as High due to lack of a control group

<sup>&</sup>lt;sup>d</sup>Overall risk of bias is Low if all domains are rated as low, High if at least one domain is assessed as high, and Unclear if at least one domain is assessed as unclear and no domains are assessed as high

First Author, Year		Selection		Outcome				
	Adequacy of	Representative-	Total	Assessment of	Same method of	Response rate	Total	/5
	case definition	ness of the sample	/2	outcome	ascertainment for	/1	/3	
	/1	/1		/1	entire sample			
					/1			
Aziz, 2004	0	0	0	0	1	0	1	1
Beaujouan, 2005	1	0	1	0	1	0	1	2
Chang, 2013	1	0	1	0	1	1	2	3
Chen, 2008	1	0	1	0	1	0	1	2
Doppia, 2011	1	1	2	0	1	1	2	4
Elovaino, 2015	1	1	2	0	1	1	2	4
Gander, 2000	1	1	2	0	1	1	2	4
Harbeck, 2015	1	0	1	0	1	1	2	3
Heponiemi, 2014	1	1	2	0	1	1	2	4
Jackson, 2017	0	0	0	0	1	0	1	1
Kanieta, 2011	1	0	1	0	1	1	2	3
Lindfors, 2006	1	1	2	0		1	2	4
Mahmood, 2017	1	0	1	0	1	0	1	2
Nishimura, 2014	1	1	2	0	1	0	1	3
Pit, 2014	1	0	1	0	1	1	2	3
Pit, 2016	1	0	1	0	1	1	2	3
Roberts, 2014	1	1	2	0	1	0	1	3
Saadat, 2016	1	1	2	0	1	1	2	4
Saadat, 2017	1	1	2	0	1	1	2	4
Sanches, 2015	1	0	1	0	1	0	1	2
Sargent, 2009	1	0	1	0	1	0	1	2

First Author, Year		Selection		Outcome				
	Adequacy of	Representative-	Total	Assessment of	Same method of	Response rate	Total	/5
	case definition	ness of the sample	/2	outcome	ascertainment for	/1	/3	
	/1	/1		/1	entire sample			
					/1			
Sende, 2010	1	0	1	0	1	0	1	2
Sexton, 2001	1	0	1	0	1	0	1	2
Shanafelt, 2005	1	0	1	0	1	1	2	3
Shanafelt, 2010	1	1	2	0	1	0	1	3
Shanafelt, 2014	1	0	1	0	1	1	2	3
Shirom, 2006	1	1	2	0	1	1	2	4
Shirom, 2010	1	1	2	0	1	1	2	4
Smith, 2016	1	0	1	0	1	1	2	3
Starmer, 2016	1	1	2	0	1	1	2	4
Tanti, 2017	1	0	1	0	1	0	1	2
Tokuda, 2009	1	1	2	0	1	1	2	4
Vela-Bueno, 2008	1	1	2	0	1	1	2	4
Wada, 2010	1	1	2	0	1	0	1	3

<sup>&</sup>lt;sup>a</sup>Assessed using the Newcastle-Ottawa Quality Assessment Scale, adapted for cross-sectional studies

<sup>&</sup>lt;sup>b</sup>An overall score of 4 to 5 stars is considered as low risk of bias, 3 as unclear risk of bias, and 2 or less as high risk of bias. For response rate, ≥50% was used as the criterion to be awarded a star

Summary of quality assessments for non-comparative studies (n=1)<sup>a</sup>

First Author,		Selection		Exposu	re		Outcome			Total
Year	Adequacy of case definition /1	Representat- iveness of the sample /1	Total /2	Ascertain- ment of exposure	Total /1	Assessment of outcome /1	Same method of assessment for entire sample	Loss to follow-up /1	Total /3	Score <sup>b</sup> /6
Gander, 2008	1	1	2	0	0	0	<b>/1</b> 1	1	2	4

<sup>&</sup>lt;sup>a</sup>Assessed using the Newcastle-Ottawa Quality Assessment Scale, adapted by the authors to be suitable to the non-comparative design

<sup>&</sup>lt;sup>b</sup>An overall score of 5 to 6 stars is considered as low risk of bias, 3 to 4 as unclear risk of bias, and 2 or less as high risk of bias 

## Supplementary file 4. Detailed study outcomes

## Physician health and wellness outcomes and associations with fatigue

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	
		time points		time points	
Surgeons					
Jackson, 2017	CS	Not feeling well rested: self-	71% healthy, 28% unhealthy in	Job satisfaction: Abridged	Job satisfaction in those more vs. less satisfied:
		reported as 'unhealthy'	terms of being well rested	Job in General Scale;	Healthy (well rested): 85% vs. 58%, p<0001;
RoB: high				grouped into more or less	Unhealthy (not well rested): 15% vs. 42%, p<0.001
		Time points NR		satisfied using the median	
				Time points NR	
Nishimura, 2014	CS	Sleep hours/night: self-	Mean±SD sleep: 5.94±1.08h	Burnout: Japanese MBI	1) Mean±SD sleep for not burned out vs. mild to
		reported (continuous)		(severe: EE >4.0 and either	moderate vs. severe: 6.07±1.15 vs. 5.88±0.94 vs.
RoB: unclear				DP >2.6 or PE <4.17)	5.63±0.94, p<0.05;
		Time points NR			2) Association between sleep and burnout (OR
				Time points NR	(95% CI)): bivariate 0.67 (0.61-0.73), p<0.001;
					multivariate including work characteristics and
				<b>10</b> ,	mental health: 0.84 (0.75-0.94), p=0.002.
Sargent, 2009	CS	Sleep deprivation: self-	21% none, 48% a little, 23%	Burnout: MBI (norms NR);	1) Positive correlation between sleep deprivation
		reported on a 4-point scale	quite a bit, 8% a lot	Marital satisfaction: RDAS;	and EE, DP, psychological distress, lower marital
RoB: high		(none, a little, quite a bit, a		Psychological morbidity:	satisfaction, all p<0.001. No relationship with PA.
		lot)		GHQ-12 score ≥4	
		Time points NR		Time points NR	
<b>Anesthesiologists</b> <sup>a</sup>					
Lederer, 2006	ВА	24-h shift with on-call duty;	Mean±SD sleep: 4.1±1.7h;	Stress during duty: 4-point	1) Mean stress score during duty: 2.1.
		Sleep hours and	Number of interruptions:	scale from 'calm' to 'very	
RoB: high		interruptions: self-reported;	0.8±1.1;	demanding'	
		Tiredness: VAS from 0 (low)	Tiredness pre- vs. post-duty:		
		to 100 (high)	30.9±27.5 vs. 59.5±18.9,	Assessed post-duty	
			p=0.01.		
		Assessed pre- and post-duty			

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome		
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and			
		time points		time points			
Leitchtfried, 2011	TS	24-h shift;	ESS (mean (range)): 7.4 (4-12);	aMT6-s: urinalysis	1) aMT6-s over shift, mean (95% CI): higher at		
		Sleepiness: ESS (range: 0-	Mean±SD sleep hours:		11:00AM pre- (12.2 (6.3-8.1)) and post-shift (9.3		
RoB: high		24);	1) pre-study: 7.74±1.35h;	Assessed at 4-h intervals	(3.7-14.9)) vs. during, p=0.016;		
		Sleep hours: self-reported	2) Pre-24-h shift (11h00 on day	from 07:00 to 11:00	2) Correlations between sleep and aMT6-s (data		
		(continuous)	1: 0.13±0.35h, 19:00 on day 1:		NR): mild for sleep duration the night prior with		
			6.99±0.68h);		aMT6-s at 3PM the following day; sleep on night 2		
		Sleepiness assessed pre-shit,	3) During the 24-h shift (07h00		with aMT6-s at 3PM the next day; total sleep with		
		sleep hours pre, during and	on day 2: 0.0±0.0h, 19h00 on		aMT6-s at 11AM on third day; moderate for sleep		
		post-shift	day 2, 5.49±1.95h);		on first night with aMT6-s at 7AM and 11AM pre-		
			4) Post-24-h shift (11h00 on		shift, 11PM during 24-h shift and 11AM post-shift;		
			day 3: 0.5±0.71h, 19h00 on		total sleep pre-shift and nocturnal sleep during 24-		
			day 3: 7.06±1.18h).		h shift with aMT6-s at 11PM during shift; total		
					sleep with aMT6-s at 3PM on first and second day,		
					11PM on second day;		
					3) Correlations between ESS and aMT6-s:		
					moderate for aMT6-s at 7AM during shift, 11AM		
					on day off.		
Beaujouan, 2005	CS	Sleep deprivation: 4-point	48.8% always or frequently	Substance abuse: 93-item	1) 60.6% with drug dependence vs. 46.0% of those		
		scale (always, frequently,	feel sleep deprived	addiction and substance	without reported sleep difficulties, p<0.001.		
RoB: high		rarely, never)		abuse questionnaire	2) OR (95% CI) of addiction for frequently/always		
					vs. rarely/never sleep deprived: tobacco 1.42		
		Time points NR		Time points NR	(1.04-1.94); tranquilizer/hypnotics 3.26 (2.12-		
					5.02).		
Doppia, 2011	CS	Insufficient sleep: 4-point	28.9% reported insufficient	Burnout: CBI (mild: 1-2.4,	1) Frequency of burnout by response for sleep		
		scale (no, not really, sort of,	sleep during work time	moderate: 2.5-3.5, severe:	sufficiency: 47.6% for no/not really, 16.3% for sort		
RoB: low		yes)		3.6-5)	of/yes, p<0.001.		
		Time points NR		Time points NR			
Lindfors, 2006	CS	Sleep hours/day: self-	Sleep hours (mean (range)): 7	Stress: MOSQ on a 3-point	1) Sleep sufficiency predicted stress symptoms:		
		reported to the nearest	(5-9)	scale (no, to some extent,	bivariate $\beta$ =-0.362, p<0.001; multivariate including		
RoB: low		0.5h;		clearly);	gender, sick leave, suicide $\beta$ =-0.269, p<0.001;		
		Adequacy of sleep and rest:		Thoughts of suicide: 4-point	2) Sleep disturbance associated with thoughts of		
		self-reported (yes/no)		scale ('never' to 'have tried')	suicide, p=0.009.		

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome	
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and		
		time points		time points		
		Time points NR		Time points NR		
Saadat, 2015	CS	Sleep deprivation (<7h/24-h)	·	Simple cognitive tests: VAS	Regular day v. post-call day, mean±SD scores:	
		due to 17-h overnight shift;	regular day vs. post-call day:	from 0 (not at all) to 100	1) Simple cognitive tests: energetic 6.04±2.27 vs.	
RoB: low		Sleepiness and alertness:	2.99±2.18 vs. 6.79±2.30,	(extremely);	2.53±1.87, confident 7.03±1.83 vs. 4.98±2.29,	
		VAS from 0 (not at all) to	p<0.001	Mood disturbance: PMS	irritable 2.03±1.94 vs. 4.86±2.16, sleepy 2.99±2.1	
		100 (extremely)		(scoring NR)	vs. 6.79±2.30, talkative 4.46±1.74 vs. 2.41±1.97, a	
					p<0.001; jittery 1.44±1.74 vs. 3.12±2.34, p=0.003;	
		All assessed on a regular day		All assessed on a regular day	anxiousness ns;	
		and a post-call day		and a post-call day	2) PMS: tension 13.48±2.71 vs. 15.43±4.46,	
					p=0.049; anger 15.24±4.41 vs. 18.14±5.92,	
					p=0.005; fatigue 10.14±2.63 vs. 20.05±6.87,	
					p<0.001; confusion 10.57±1.69 vs. 12.57±4.24,	
					p=0.025; vigor 24.05±6.75 vs.16.67±5.70, p<0.001	
					depression: ns; total mood disturbance:	
					42.57±15.26 vs. 70.90±6.91, p<0.001.	
ER or ICU physicia	15					
Dutheil, 2013	RCT	14-h or 24-h shift;	1) Sleep duration and quality	Stress: VAS from 0 (low) to	1) Stress: higher following 14-h and 24-h shifts vs.	
		Sleep hours: self-reported	lower during shifts (14h and	100 (high);	the control day, p<0.05 (data NR);	
RoB: high		sleep and wake time;	24h) than any other day, and	IL-8: urinalysis	2) IL-8: higher following 24-h shift vs. control	
. 0		•				
. 0		Sleep quality: VAS from 1	lower during the 24-h vs. 14-h		(p=0.007) and 14-h shift (p=0.015); ns difference	
· U		Sleep quality: VAS from 1 (low) to 100 (high);	lower during the 24-h vs. 14-h shift (p<0.05);	Assessed at 08:30 and 18:30	(p=0.007) and 14-h shift (p=0.015); ns difference between 14-h shift and control day;	
0		Sleep quality: VAS from 1 (low) to 100 (high); Mental and physical fatigue:	lower during the 24-h vs. 14-h shift (p<0.05); 2) Mental and physical fatigue	Assessed at 08:30 and 18:30 on each day of protocol	(p=0.007) and 14-h shift (p=0.015); ns difference between 14-h shift and control day; 3) Correlations with IL-8: sleep hours pre-24-h	
		Sleep quality: VAS from 1 (low) to 100 (high); Mental and physical fatigue: VAS from 1 (low) to 100	lower during the 24-h vs. 14-h shift (p<0.05); 2) Mental and physical fatigue higher after 14-h and 24-h shift		(p=0.007) and 14-h shift (p=0.015); ns difference between 14-h shift and control day; 3) Correlations with IL-8: sleep hours pre-24-h shift, r=-0.627, p=0.007; poor sleep quality during	
		Sleep quality: VAS from 1 (low) to 100 (high); Mental and physical fatigue:	lower during the 24-h vs. 14-h shift (p<0.05); 2) Mental and physical fatigue		(p=0.007) and 14-h shift (p=0.015); ns difference between 14-h shift and control day; 3) Correlations with IL-8: sleep hours pre-24-h shift, r=-0.627, p=0.007; poor sleep quality during 14-h and 24-h shifts, r=0.452, p=0.031;	
		Sleep quality: VAS from 1 (low) to 100 (high); Mental and physical fatigue: VAS from 1 (low) to 100 (high)	lower during the 24-h vs. 14-h shift (p<0.05); 2) Mental and physical fatigue higher after 14-h and 24-h shift		(p=0.007) and 14-h shift (p=0.015); ns difference between 14-h shift and control day; 3) Correlations with IL-8: sleep hours pre-24-h shift, r=-0.627, p=0.007; poor sleep quality during 14-h and 24-h shifts, r=0.452, p=0.031; 4) Multivariable regression: 24-h shift increased II	
		Sleep quality: VAS from 1 (low) to 100 (high); Mental and physical fatigue: VAS from 1 (low) to 100 (high) Assessed on day prior to	lower during the 24-h vs. 14-h shift (p<0.05); 2) Mental and physical fatigue higher after 14-h and 24-h shift		(p=0.007) and 14-h shift (p=0.015); ns difference between 14-h shift and control day; 3) Correlations with IL-8: sleep hours pre-24-h shift, r=-0.627, p=0.007; poor sleep quality during 14-h and 24-h shifts, r=0.452, p=0.031; 4) Multivariable regression: 24-h shift increased IL 8 by 1.9ng vs. control day, p=0.007; ns association	
		Sleep quality: VAS from 1 (low) to 100 (high); Mental and physical fatigue: VAS from 1 (low) to 100 (high)  Assessed on day prior to shift; during shift; each day	lower during the 24-h vs. 14-h shift (p<0.05); 2) Mental and physical fatigue higher after 14-h and 24-h shift		(p=0.007) and 14-h shift (p=0.015); ns difference between 14-h shift and control day; 3) Correlations with IL-8: sleep hours pre-24-h shift, r=-0.627, p=0.007; poor sleep quality during 14-h and 24-h shifts, r=0.452, p=0.031; 4) Multivariable regression: 24-h shift increased IL 8 by 1.9ng vs. control day, p=0.007; ns association with 14-h shift, mental or physical fatigue, sleep	
		Sleep quality: VAS from 1 (low) to 100 (high); Mental and physical fatigue: VAS from 1 (low) to 100 (high)  Assessed on day prior to shift; during shift; each day of protocol (work, off,	lower during the 24-h vs. 14-h shift (p<0.05); 2) Mental and physical fatigue higher after 14-h and 24-h shift		(p=0.007) and 14-h shift (p=0.015); ns difference between 14-h shift and control day; 3) Correlations with IL-8: sleep hours pre-24-h shift, r=-0.627, p=0.007; poor sleep quality during 14-h and 24-h shifts, r=0.452, p=0.031; 4) Multivariable regression: 24-h shift increased IL 8 by 1.9ng vs. control day, p=0.007; ns association	
		Sleep quality: VAS from 1 (low) to 100 (high); Mental and physical fatigue: VAS from 1 (low) to 100 (high)  Assessed on day prior to shift; during shift; each day of protocol (work, off, clerical, control)	lower during the 24-h vs. 14-h shift (p<0.05); 2) Mental and physical fatigue higher after 14-h and 24-h shift		(p=0.007) and 14-h shift (p=0.015); ns difference between 14-h shift and control day; 3) Correlations with IL-8: sleep hours pre-24-h shift, r=-0.627, p=0.007; poor sleep quality during 14-h and 24-h shifts, r=0.452, p=0.031; 4) Multivariable regression: 24-h shift increased IL 8 by 1.9ng vs. control day, p=0.007; ns association with 14-h shift, mental or physical fatigue, sleep	
Sende, 2012	CS	Sleep quality: VAS from 1 (low) to 100 (high); Mental and physical fatigue: VAS from 1 (low) to 100 (high)  Assessed on day prior to shift; during shift; each day of protocol (work, off,	lower during the 24-h vs. 14-h shift (p<0.05); 2) Mental and physical fatigue higher after 14-h and 24-h shift		(p=0.007) and 14-h shift (p=0.015); ns difference between 14-h shift and control day; 3) Correlations with IL-8: sleep hours pre-24-h shift, r=-0.627, p=0.007; poor sleep quality during 14-h and 24-h shifts, r=0.452, p=0.031; 4) Multivariable regression: 24-h shift increased IL 8 by 1.9ng vs. control day, p=0.007; ns association with 14-h shift, mental or physical fatigue, sleep	
	CS	Sleep quality: VAS from 1 (low) to 100 (high); Mental and physical fatigue: VAS from 1 (low) to 100 (high)  Assessed on day prior to shift; during shift; each day of protocol (work, off, clerical, control)	lower during the 24-h vs. 14-h shift (p<0.05); 2) Mental and physical fatigue higher after 14-h and 24-h shift vs. control day (data NR).	on each day of protocol	(p=0.007) and 14-h shift (p=0.015); ns difference between 14-h shift and control day; 3) Correlations with IL-8: sleep hours pre-24-h shift, r=-0.627, p=0.007; poor sleep quality during 14-h and 24-h shifts, r=0.452, p=0.031; 4) Multivariable regression: 24-h shift increased IL 8 by 1.9ng vs. control day, p=0.007; ns association with 14-h shift, mental or physical fatigue, sleep deprivation, 14-h shift.	

Study Risk of Bias (RoB)	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
	design	sign Assessment measure and	Baseline	Assessment measure and	
		time points		time points	
				related, organizational,	
		Time points NR		individual)	
				Time points NR	
Generalists <sup>b</sup>					
	CC	24 have an apliability with	1) Class basses are a resumed day.	Diaghamical /labamatam.	Defense a negresal shifts on effect accomplished call shifts
Harbeck, 2015	CS	24-hours on-call shift with sleep disturbance: self-	1) Sleep hours on a normal day vs. following a 24-h shift:	Biochemical (laboratory values) and physiological	Before a normal shift vs. after overnight call shift:  1) Biochemical parameters: no changes in any
RoB: unclear		reported number of sleep	<2 hours: 0 vs. 5.9%; 2-4	(heart rate variability, skin	parameter except for thyroid stimulating hormon
NOD. uncical		disturbances and hours of	hours: 5.9% vs. 47.1%; 4-6	resistance, blood pressure)	which was higher after the on-call shift ( $p = 0.049$ )
		sleep per night	hours: 11.8% vs. 35.3%; >6	stress parameters	data NR);
		sicep per mgm	hours: 82.4% vs. 11.8%	otress parameters	<ul><li>2) Physiological parameters: no significant change</li></ul>
		Assessed before a normal	2) Number of sleep	Assessed before a normal	in any parameter
		day shift, and after a 24-h on	disturbances a normal day vs.	day shift, and after a 24-h	
		call shift	following a 24-h shift:	on call shift	
			0: 82.4% vs. 11.8%; 1: 11.8%		
			vs. 35.3%; 2: 5.9% vs. 47.1%; 3:		
			0% vs. 5.9%; 4: 0% vs. 0%; >4:		
			0% vs. 0%		
Pit, 2014	CS	Work-related sleep	Work-related sleep	Early retirement (<65 years)	For sleep disturbance a few times a year to every
		disturbance: 7-point scale	disturbance: 41% never, 59% a	intentions (yes/no)	day vs. never:
RoB: unclear		from 'never' to 'every day'	few times a year to every day	T' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' ' '	1) Intention to retire early: 74% vs. 26%, p<0.01;
		Time and into NID		Time points NR	2) Association with intention to retire early (OR
		Time points NR			(95% CI)): univariate 3.6 (1.47-8.80), p<0.01;
					multivariate including work, occupational, individual factors 2.91 (1.11-7.6), p<0.05;
					4) RR (95% CI) for intention to retire early: 2.0
					(1.18-3.49); attributable fraction: 50.0%;
					population attributable fraction: 37.1%.
Pit, 2016	CS	Work-related sleep	Work-related sleep	Sickness presenteeism: 'yes'	For sleep disturbance a few times a year to every
		disturbance: 7-point scale	disturbance: 41% never, 59% a	response indicated 1 or	day vs. never:
RoB: unclear		from 'never' to 'every day'	few times a year to every day	more days	1) Sickness presenteeism: 32% vs. 68%, p=0.018;

Study Risk of Bias (RoB)	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
	design	Assessment measure and time points	Baseline	Assessment measure and	
				time points	
		Time points NR		Assessed for the past 12	2) Association with sickness presenteeism (OR
				months	(95% CI)): 2.92 (1.19-7.16), p=0.02.
Roberts, 2014	CS	Fatigue: LAS from 0 (low) to	Mean (SD) score: 5.8 (2.4) for	Impact of fatigue on daily	1) 8.7% of hospitalists and 4.3% of outpatient
		10 (high)	hospitalists; 5.9 (2.4) for	activities (falling asleep	general internists had fallen asleep while driving
RoB: unclear			general internists	while driving) (yes/no)	due to fatigue.
		Assessed for the past week			
				Time points NR	
Vela-Bueno, 2008	CS	Sleep Quality: PSQI	Prevalence (% (95% CI)):	Burnout: PBM with a 7-point	Low vs. high burnout, mean±SD:
		(Spanish): score ≥5 indicates	1) Sleep-onset latency >30	scale from 1 (never) to 7	1) Global PSQI: 2.72±2.22 vs. 7.24±4.17, p<0.001;
RoB: low		low quality (range; 0 to 21);	minutes: 8.4 (4.8-11.9);	(always)	2) PSQI subscores: sleep quality: 0.54±0.57 vs.
		Insomnia: DSM-IV criteria	2) Wake time after sleep onset		1.40±0.83, p<0.001; sleep latency: 0.51±0.80 vs.
			>30 minutes: 15.4 (10.8-19.9);	Time points NR	1.38±1.03, p=0.002; sleep duration: 0.45±0.64 vs.
		Time points NR; insomnia	3) Early morning awakening:		1.16±0.92, p=0.003; sleep efficiency: 0.21±0.57 vs
		symptoms in past month	22.5 (19.5-30.4);		0.77±0.98, p=0.018; sleep disturbance: ns; use of
			4) Nonrestorative sleep: 22.5		medication: 0.14±0.49 vs. 0.57±0.83, p=0.032;
			(17.2-27.7);		daytime dysfunction: 0.52±0.73 vs. 1.57±0.88,
			5) Daytime impairment for ≥5		p=0.002.
			days in past month: 14.2 (9.7-		3) Prevalence (95% CI) of insomnia symptoms:
			18.6);		sleep latency: 5.5% (2.5-11.5%) vs. 21.1% (10.5-
			6) Insomnia: 18.8 (13.8-23.7).		31.6%), p=0.015; wake time >30 min after sleep
					onset: 9.4% (1.6-17.1%) vs. 25.5% (14.2-37.7%),
					p=0.029; early awakening: 14.5% (5.1-23.8%) vs.
					45.6 (32.7-58.4%), p<0.001; somewhat/very
					dissatisfied with sleep: 5.5% (2.5-11.5%) vs. 50%
					(37.1-62.8%), p<0.001; day impairment: 5.5% (2.5
					11.5%) vs. 38.2% (25.6-50.7%), p<0.001; insomnia
					7.3% (0.4-14%) vs. 39.7% (27.1-52.2%), p<0.001.
Oncologists					
Shanafelt, 2005	CS	Fatigue: LASA QOL ≤7;	75% had a high level of fatigue;	Wellbeing: 10-item LASA	1) Sleep deprivation for high vs. low overall well-
		Sleep deprivation: 10-point	Mean±SD sleep score:	QOL, high ≥8 vs. low ≤7	being (mean±SD): 3.9±2.57 vs. 5.1±2.60, p=0.0004
RoB: unclear		Likert scale from 0 (not at	4.5±2.65.		2) Lower fatigue predicted overall wellbeing in a
		all) to 10 (stressful as can		Time points NR	multivariate model including personal and
		be)			professional characteristics, p=0.002.

•	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
	design	Assessment measure and	Baseline	Assessment measure and	
		time points		time points	
		Time points NR			
Shanafelt, 2014	CS	Fatigue: 10-point LAS (lower	Mean±SD fatigue score:	Satisfaction with WLB: 5-	1) OR (95%CI) of lower satisfaction predicted by
,		scores indicate greater	5.7±2.4	point Likert scale from	high fatigue (vs. not) in multivariate model
RoB: unclear		fatigue)		'strongly agree' to 'strongly	including personal and work-related factors, and
				disagree'	burnout: 0.489 (0.337-0.710), p<0.001.
		Time points NR		-	
		UA		Time points NR	
Mixed groups of pl	nysicians				
Aziz, 2004	CS	Working while fatigued: 5-	NR	Stress: 47-item	1) Sources of stress: working while fatigued had a
		point scale from 'extreme'		questionnaire with a 5-point	mean±SD score of 2.44±1.20, factor loading:
RoB: high		to 'a little'		scale from 'extreme' to 'a	0.653, in factor analysis;
				little'	2) Inverse correlation between stress and working
		Time points NR			while fatigued: r=-0.270 (significance level NR).
			<b>/ (</b> )	Time points NR	
Chen, 2008	CS	Sleepiness: ESS score ≥11	Mean±SD ESS score: 7.8±4.0,	Impact on work and	1) Impact score correlated with ESS, r=0.31,
			range: 0-20, 23% had scores	personal life: Impact	p<0.05;
RoB: high		Time points NR	≥11.	Questionnaire with a 5-	2) ESS score was higher among physicians who
				point Likert scale from 1	agree/strongly agree vs. other response: worried
				(strongly agree) to 5	about having a car accident while driving home
				(strongly disagree)	post-call: 5.4 vs. 7.0, p<0.001; sleep loss has a
					major impact on personal life: 8.4 vs. 7.0, p=0.01;
				Time points NR	3) Higher ESS scores predicted by impact score in
					multivariate regression including personal and
51 : 2045		Cl	14		work-related factors: β=0.11, p=0.005.
Elovaino, 2015	CS	Sleeping problems: Jenkins	Mean±SD score:	Jobs demands: 5 items	There was no association between sleeping
DaD. la		Scale with a 6-point scale	2006: 2.30 (1.00);	scored on a 5-point scale	problems in 2006 and job demands or control in
RoB: low		from 1 (never) to 6 (every	2010: 2.35 (1.05).	from 1 (strongly disagree) to	2010.
		night)		5 (strongly agree);	
		Accorded in 2006 and 2010		Job control: 3 items derived	
		Assessed in 2006 and 2010		from the Karasek Job	
				Questionnaire	

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and time points	Baseline Asse	Assessment measure and	
				time points	
Heponiemi, 2014	CS	Sleeping problems: Jenkins	Mean±SD (range) score:	Psychological distress: GHQ-	1) Sleeping problems associated with job
		Scale <sup>81</sup> with a 6-point scale	2.30±1.00 (1-6)	12 with a 4-point scale (low	satisfaction, β=-0.12, p<0.001, psychological
RoB: low		from 1 (never) to 6 (every		to high);	distress, β=0.18, p<0.001;
		night)		Job satisfaction: JDS with a	2) Total indirect effect of on-call duty through two
				Likert scale from 1 (strongly	mediators (sleeping problems, work interference
		Assessed in 2006		disagree) to 5 (strongly	with family) (R <sup>2</sup> (95% CI)): job satisfaction 0.06 (-
				agree)	0.059, -0.016), p<0.001; psychological distress 0.16
					(0.023, 0.081), p<0.001.
				Assessed in 2010	
Mahmood, 2016	CS	Sleep deprivation: self-	Mean±SD hours:	Alcohol use disorders:	There was no association between hours of sleep
		reported mean hours of	4 years: 4.52 (2.79);	Modified 9-item version of	when on call and hazardous drinking behaviours
RoB: high		sleep when on call	10 years: 5.38 (6.36);	the Alcohol Use Disorder	(p=0.732)
			15 years: 6.41 (7.14).	Identification Test (AUDIT)	
		Assessed at 4 years, 10		≥6 for men and ≥5 for	
		years, and 15 years post-		women.	
		graduation			
				Assessed at 4 years, 10	
				years, and 15 years post-	
				graduation	
Shirom, 2010	CS	Tiredness and exhaustion:	NR	Burnout: SMBM on a 7-	1) Correlation between physical fatigue subscale
		SMBM Physician Fatigue		point scale from 1 (almost	and overall burnout: 0.88, p<0.05;
RoB: low		Subscale on a 7-point scale		never) to 7 (always)	2) In a predictive structural model for burnout,
		from 1 (almost never) to 7			physical fatigue accounted for unique variance in
		(always)			the burnout items, not accounted for by total burnout (R <sup>2</sup> =0.24).
		Time points NR			
Smith, 2017	CS	Sleep deprivation: self-	NR	Mental and physical illness:	Some physicians reported developing mental
		reported via open-ended		self-reported via open-	illness (e.g., bipolar disorder, alcohol misuse) due
RoB: unclear		comments		ended comments	to tiredness and stress at work; others developed
					physical health problems due to sleep deprivation
		Time points NR		Time points NR	poor eating habits and lack of exercise.

Study	Study	tudy Exposures or interventions C		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	_
		time points		time points	
Starmer, 2016	CS	Sleep deprivation: <7 hours	27.7% sleep deprived	Burnout, satisfaction with	≥7-h vs. <7-h sleep:
		sleep in a typical 24-h period		career and life, balanced	1) Burnout (% strongly agree/agree): 26.4% vs.
RoB: low		(self-reported)		personal and professional	39.6%, p<0.05; career satisfaction (% strongly
				commitments: Each on a 5-	agree/agree): ns; life satisfaction (%
		Time points NR		point Likert scale (strongly	completely/very satisfied): 76.4% vs. 55.9%,
				agree to strongly disagree)	p<0.05; balanced personal and professional
					commitments (% completely/very satisfied): 49.79
				Time points NR	vs. 26.1%.
					2) <7-h sleep (vs. ≥7-h) (OR, 95% CI) associated
					with life satisfaction 0.44 (0.29-0.67), p<0.05;
					balanced personal/professional commitments 0.4
					(0.31-0.71), p≤0.05, in a model including work and
					personal factors.
Tokuda, 2009	CS	Sleep hours/day: self-	Mean±SD (range) sleep	Burnout: MBI (Japanese)	Maximum likelihood estimates±SE:
		reported (continuous)	hours/day: 6±0.9 (3-8)	with a 7-point Likert	1) Sleeping time to job satisfaction: group
RoB: low				scale: 0 (none) to 6 (every	0.990±0.458, p=0.031; ns for men; women
		Time points NR (included		day);	1.711±0.805, p=0.034;
		weekday and weekends)		Job satisfaction: JHPSS	2) Sleeping time to EE: group -0.219 ±0.070,
				with a 5-point Likert	p=0.002; men -0.215±0.082, p=0.009; ns for
				scale: 1 (strongly	women.
				disagree) to 5 (strongly	
				agree)	
				Time points NR	
Wada, 2010	CS	Sleep hours/day: Self-	<5 hours: 8.7% men, 9.9%	Depression: QIDS-SR;	1) Sleep hours for those with vs. without
		reported (continuous)	women; 5 to <6 hours: 32.3%	Japanese score <5 (no	depressive symptoms: <5: 18.7% vs. 7.7% men,
RoB: unclear			men, 34.6% women; 6 to <7	symptoms) to >20 (very	20.5% vs. 8.7% women; 5 to <6: 33.7% vs. 32.2%
		Assessed for past month	hours: 46.0% men, 43.7%	severe symptoms)	men, 38.6% vs. 34.2% women; 6 to <7: 35.1% vs.
		when not completing	women; ≥7 hours: 13.0% men,		46.9% men; 31.8% vs. 45.1% women;
		overnight work	11.8% women.	Assessed for past 7 days	2) Association between <5h sleep (vs. 6-7h) and depressive symptoms (OR (95% CI)): univariate
					2.79 (1.96-3.95) for men, 2.65 (1.47-4.78) for
					women; multivariate (including age and workload

Study	Study	Exposures or interventions		Outcomes Associations betw	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	
		time points		time points	
					factors) 2.70 (1.82-4.03) for men, 2.38 (1.11-5.10)
					for women.

<sup>&</sup>lt;sup>a</sup>Includes studies of anesthetists, where these were physicians.

AM: morning; aMT6-s: melatonin metabolite; BA: before-after; CI: confidence interval; CBI: Copenhagen Burnout Inventory; CS: cross-sectional; DP: depersonalization; DSM: Diagnostic and Statistical Manual of Mental Disorders; EE: emotional exhaustion; ER: emergency; ESS: Epworth Sleepiness Scale; GHQ: General Health Questionnaire; h: hour(s); ICU: intensive care unit; IL-8: interleukin-8; JDS: Job Diagnostic Survey; JHPSS: Japanese Hospital Physicians Satisfaction Scale; LAS: linear analog scale; LASA: linear analog assessment scales; MBI: Maslach Burnout Inventory; MOSQ: Modified Occupational Stress Questionnaire; min: minute(s); NA: not applicable; NR: not reported; ns: not statistically significant; OR: odds ratio; PA: personal achievement; PBM: Pines Burnout Measure; PE: professional efficacy; PM: afternoon; PMS: Profile of Mood States; PSQI: Pittsburgh Sleep Quality Index; QIDS-SR: Quick Inventory Depressive Scale - Self-Reported; QOL: Quality of Life; RCT: randomized controlled trial; RDAS: Revised Dyadic Adjustment Scale; RoB: Risk of Bias; SD: standard deviation; SE: standard error; SMBM: Shirom-Melamed Burnout Measure; TS: time series; US: United States of America; VAS: visual analog scale; vs.: versus; WLB: work-life balance

<sup>&</sup>lt;sup>b</sup>Includes primary care physicians, internal medicine physicians, and general practitioners.

Performance and safety outcomes related to fatigue or sleep loss among physicians in independent practice

Study	Study	Exposures or intervention		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	ign Assessment measure and	Baseline	Assessment measure and	
		time points	time points		
Surgeons					
Uchal, 2005	RCT	Sleep deprivation from a 24-	Median (range) sleep hours:	Surgical performance:	Post call vs. post-work:
		h call shift vs. 8-h work;	1.5 (0-3) post-call vs. 6.5 (5-	laparoscopic surgical	1) Product quality: no difference in accuracy
RoB: unclear		Sleep hours: self-reported	9) post-work, p<0.05;	simulator(Minimally Invasivs	error, tissue damage, leak rate;
		(continuous);	Median ESS score: 7.0 post-	Surgical Trainer-Virtual	2) Procedure effectiveness: no difference in
		Sleepiness: ESS (moderate:	call vs. 5.5 post-work, ns.	Reality) for product quality,	goal-directed actions, non-goal directed
		10-15, severe: ≥16)		procedure effectiveness	actions, operating time.
		Assessed post-call and post-		Assessed post-call and post-	
		work		work	
Chu, 2011	СО	Sleep deprivation: self-	Of 4,047 procedures, 83	Surgical performance: CABG,	For 0-3 vs. 3-6 vs. >6 hours of sleep: no
		reported hours, moderate	(2.1%) performed by	ACC	difference in CABG or ACC.
RoB: low		(3-6h) or severe (<3h)	severely sleep-deprived and		
			1,595 (39.4%) moderately	Assessed during surgery	
		Assessed the night before	sleep-deprived surgeons		
		surgery			
Ellman, 2004	СО	Sleep deprivation:	Of 6,751 procedures, 339	Surgical performance: CABG,	Sleep deprived vs. non-sleep deprived: no
		performed a case starting	(5%) performed by sleep-	ACC	difference in CABG or ACC.
RoB: low		22:00 to 05:00, or ending	deprived surgeons		
		22:00 to 07:30 and another		Assessed during surgery	
		case in the next 24-h			
Govindarajan,	СО	Sleep deprivation: treated	NR	Surgical performance:	Sleep deprived vs. non-sleep deprived: no
2015		patients from midnight to		duration of surgery	difference in duration of surgery, even after
		07:00 and performed a			stratification by type of procedure.
RoB: low		subsequent case on the			
		same day			
Amirian, 2014	ВА	17-h night call shift;	Naps pre-call: 11 (37%)	Surgical performance:	Pre- vs. post-call:
		Sleep hours during the shift:	napped for median (IQR) 90	LapSimGyn laparoscopic	1) LapSimGyn: no difference in total time,
RoB: high		Wrist-mounted Micro-Mini-	(58-128) min;	simulation for time, blood	blood loss, instrument path length, instrument
		Motionlogger;	Median (IQR) sleep: 91 (62-	loss, instrument path;	angular path; napping did not affect
		Sleepiness: KSS	123) min on the pre-call night vs. 430 (329-449) on	D2 test of attention and concentration	performance;

Study	Study	Exposures or intervention		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and time points	Baseline	Assessment measure and time points	•
		Assessed on pre-call and on-	the on-call night, p<0.001;		2) D2 test: improvement in concentration,
		call day; sleepiness assessed	Sleep on-call: 12 (40%) slept	Assessed on pre-call and on-	p<0.05. No changes in any other parameters;
		during shift	for median (IQR) 98 (39-	call day	3) ns difference in laparoscopic simulation time
			135) min;		in those who slept during the shift vs. not.
			Significant development of		
			sleepiness during shift		
			(p<0.001), plateau score of		
		Uh	7 at 04:00 to 08:00.		
Gerdes, 2008	BA	On-call shift;	Fatigue differential from	Psychomotor performance:	1) Pre- to post-call: decrease in all measures of
		Fatigue: questionnaire	pre- to post-call (range): 1-7	virtual ring transfer task for	psychomotor proficiency (p<0.05, data NR)
RoB: high		designed by Behrenz &	(units unclear);	gesture-level proficiency,	except elapsed time; no change in number of
		Monga, 1999;	Sleep during call (range): 1-	hand movement	psychomotor errors; increase cognitive errors
		Sleep hours: self-reported	5h	smoothness, tool movement	(p<0.05, data NR);
		(continuous)		smoothness, elapsed time	2) Cognitive errors increased exponentially as
					fatigue ratings increased (R <sup>2</sup> =0.9219) and as
		Assessed in 3 sessions pre-		Assessed in 3 sessions pre-	hours of sleep declined (R <sup>2</sup> =0.933).
		and post-call		and post-call	
Shanafelt, 2010	CS	Degree of fatigue as a	NR	Perceived recent major	1) Prevalence of perceived recent major
		contributor to errors (self-		medical errors (self-	medical error: 8.9%;
RoB: unclear		reported)		reported)	2) Of those reporting an error, 6.9% listed
					degree of fatigue as the greatest contributing
		Assessed for the past 3		Assessed for the past 3	factor.
		months		months	
Anesthesiologists					
Lederer, 2006	BA	24-h shift, on-call duty;	Mean±SD sleep: 4.1±1.7h;	Psychomotor performance:	Pre- vs. post-duty, mean±SD:
		Sleep hours and	Number of interruptions:	reaction time, critical flicker	1) Psychometric testing: recognition reaction
RoB: high		interruptions: self-reported;	0.8±1.1;	fusion, response measure,	time (ms): 439.6±50.8 vs. 480.3±58.9; motor
		Tiredness: VAS from 0 (low)	Tiredness pre- vs. post-duty:	peripheral awareness;	reaction time (ms): 252.8±39.3 vs. 465.4±65.0;
		to 100 (high)	30.9±27.5 vs. 59.5±18.9,	Concentration ability: scale	total reaction time (ms): 690.8±73.4 vs.
			p=0.01.	of 0 (low tiredness) to 100	746.5±113.7; critical flicker fusion (Hz):
		Assessed pre- and post-duty		(maximum tiredness)	29.0±2.3 vs. 28.7±3.7; response measure
					(pixels): 647.8±126.7 vs. 598.3±138.1,
				Assessed pre- and post-duty	

Study	Study design	dy Exposures or intervention		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)		Assessment measure and	Baseline	Assessment measure and	_
		time points		time points	
					peripheral awareness task recognition time:
					58.9±59.2 vs. 51.6±47.5;
					2) Concentration ability: 26.4±23.5 vs.
					56.3±23.0, p=0.007.
Chang, 2013	CS	15-h in-house overnight call;	Median (IQR) ESS: 9 (9),	Psychomotor performance:	1) Afternoon baseline vs. pre-call: no
		Sleepiness pre-call: ESS ≥9;	64% scored ≥9;	reaction time; CCPT II; N-	difference in reaction time, CCPT, N-back, of
RoB: unclear		Sleep hours: self-reported	Median (IQR) hours slept	back; HVLT (3 trials of 12	HVLT;
		(continuous)	during shift: 1 (0-3).	words)	Morning baseline vs. post-call:
					1) No change in auditory or visual reaction
		Sleepiness assessed pre-call,		Assessed at baseline and pre-	time;
		sleep hours during call		and post-call	2) CCPT (t-scores): No change in detectability,
					response style, hit reaction time,
					omissions/commissions;
					3) N-back % accuracy: no change for auditory,
					visual, or mean N-value;
					4) HVLT (t-score): mean for trials 1-3: 48.6±7.6
					vs. 41.5±9.9 (p=0.04); delayed recall: ns;
					5) No correlation between ESS scores pre-call
					or sleep during shift and any measure of
			eer te		psychomotor performance.
Gander, 2000	CS	Nights of work-related sleep	NR	Risk of fatigue-related errors:	1) Risk of fatigue-related errors increased with
		disturbance: self-reported		questionnaire modelled after	increasing nights of work-related sleep
RoB: low		(continuous)		Gravenstein et al., 1990	disturbance: RR: 1.25, 95% CI: 1.06-1.49.
		Assessed for the past 6		Assessed for the past 6	
		months		months	
Saadat, 2017	CS	Sleep deprivation due to an	NR	Reaction time: PVT	Mean (SD) reaction time was slower post-call
		overnight call shift			(297.76 (83.75)) vs. on a regular day (266.58
RoB: low				Assessed after an overnight	(38.35)), p=0.047.
				call shift and the morning of	
				a regular (non-call) day	

Study	Study design	Exposures or intervention	posures or intervention Outcome		Associations between exposure and outcome
Risk of Bias (RoB)		Assessment measure and time points	Baseline	Assessment measure and time points	
Gander, 2008	NC	Sleep loss across	≥2 hours sleep <baseline:< td=""><td>Psychomotor performance:</td><td>1) In fixed model analysis for reaction time</td></baseline:<>	Psychomotor performance:	1) In fixed model analysis for reaction time
		consecutive working days or	8% of 24-h periods that	PVT	including sleep, time since waking, work hours
RoB: unclear		on-call work: Wrist-	included day work vs. 14%		acute sleep loss associated with slower media
		mounted Actiwatch (Mini	that included day + call;	Assessed within 2 hours pre-	reaction time, $F_{(1,184)}=5.70$ , p<0.05; longer time
		Mitter, Bend, Oregon, US),	Sleep hours: mean 0.6h less	and post-call	since waking associated with poorer
		sleep and duty diary	sleep when working day		performance on the slowest 10%, $F_{(1,185)}=5.13$
			shifts (p=0.014) and 0.8h		p<0.05;
		Assessed over a 2-week	less sleep when working day		2) Reaction time across 12 consecutive work
		period including a weekend	shifts + call (p=0.013) vs. off.		days: no change in pre-duty reaction times but
		of rostered shifts or on-call			post-duty reaction times slowed linearly,
					median -3.38, p<0.001; decline in performance
					across 10 minutes became progressively
					steeper both pre- and post-duty, p=0.020.
ER or ICU physician					
Sanches, 2015	CS	Acute sleep deprivation	Non-sleep deprived vs.	Psychomotor performance	Sleep deprived group vs. non-sleep deprived,
		(<5h of night sleep after a	sleep deprived:	via Battery Test Reaction 5	mean±SD:
RoB: high		night shift of 12h)		(v1): StimulTest, InstrucTest,	1) InstrucTest: correct answers: 169.4 (16.0) vs
		Sleep hours: 7-day	PSQI >5: 0% vs. 33%, ns;	MovemTest; TP test of visual	148.3 (28.3), p=0.070; wrong answers: ns;
		Actigraphy via SenseWear®	ESS≥10: 11% vs. 67%	attention	perfection index (%): 99.6 (0.3) vs. 98.9 (1.3),
		Pro2 Armband;	Sleep time (mean±SD) in		p=0.021; response latency (sec/click): ns;
		Sleepiness: ESS;	week before tests: duration	Assessed on morning after	2) StimulTest: correct answers: 170.7 (21.9) vs
		Sleep quality: PSQI	and number of naps higher	night shift 8	145.1 (17.9), p=0.022; wrong answers: ns;
			in sleep deprived group, but		perfection index (%): ns; response latency
		Assessed the week and	diurnal sleep hours lower,		(sec/click): 1.06 (0.1) vs. 1.24 (0.1), p=0.022;
		night before the	428.6±30.1 vs. 375.8±55.9,		3) MovemTest: ns for any parameter;
		psychomotor tests	p=0.038;		4) TP: omitted symbols: 34.2±18.4 vs.
			Sleep quality (mean±SD):		62.7±44.0, p=0.034; concentration index (%):
			week before tests: 3.3±0.7		14.1±8.9 vs. 30.0±25.9, p=0.019; quality index
			vs. 2.6±0.3, p=0.013;		(%): 13.8±8.6 vs. 29.2±26.4, p=0.031;
			night before tests: 3.1±0.8		correct/wrong symbols: ns;
			vs. 1.9±1.0, p=0.020.		Correlations between sleep and tests:
					1) TP for sleep hours nights 1-6: omitted
					symbols: r=-0.686, p=0.011 for non-sleep-

Study	Study	Exposures or intervention	osures or intervention		Associations between exposure and outcome
Risk of Bias (RoB)	design	esign Assessment measure and	Baseline	Assessment measure and	
		time points		time points	
					deprived, ns for sleep-deprived; concentration
					index (%): r=-0.359, p=0.037 for sleep-
					deprived, ns for non-sleep deprived; r=-0.359,
					p=0.037 for the group; no other significant correlations;
					2) No correlation between PSQI, ESS and any or
					the psychomotor tests.
Generalists <sup>b</sup>					
Harbeck, 2015	CS	24-hours on-call shift with	1) Sleep hours on a normal	Neurocognitive parameters:	Intrinsic alertness, focused attention and
		sleep disturbance: self-	day vs. following a 24-h	computerized attentional	vigilance were similar on both occasions;
		reported number of sleep	shift: <2 hours: 0 vs. 5.9%;	test (vigilance, alertness); D2	Phasic alertness improved following the on-cal
		disturbances and hours of	2-4 hours: 5.9% vs. 47.1%;	letter cancellation test	shift: mean (SD) 24.8 (15.6) vs. 38.3 (21.5), p =
		sleep per night	4-6 hours: 11.8% vs. 35.3%;	(divided attention); Trail	0.022.
			>6 hours: 82.4% vs. 11.8%	Making Test (visual	
		Assessed before a normal	2) Number of sleep	attention, task switching);	
		day shift, and after a 24-h	disturbances a normal day	Digit Span, Digit Symbol	
		on call shift	vs. following a 24-h shift:	Substitution Test, Weschler	
			0: 82.4% vs. 11.8%; 1: 11.8%	Memory Scale (memory	
			vs. 35.3%; 2: 5.9% vs. 47.1%;	functions)	
			3: 0% vs. 5.9%; 4: 0% vs. 0%;	, and a second	
			>4: 0% vs. 0%	Assessed before a normal	
			7 1. 6/6 43. 6/6	day shift, and after a 24-h on	
				call shift	
Mixed specialties o	r undefine	ed populations		Sun Simil	
Chen, 2008	CS	Sleepiness: ESS score ≥11	Mean±SD ESS score:	Impact on work and personal	1) Impact score correlated with ESS, r=0.31,
			7.8±4.0, range: 0-20, 23%	life: Impact Questionnaire	p<0.05;
RoB: high		Time points NR	had scores ≥11.	with a 5-point Likert scale	2) ESS score was higher among physicians who
				from 1 (strongly agree) to 5	agree/strongly agree vs. other response:
				(strongly disagree)	written an incorrect order: 8.8 vs. 7.3, p=0.02;
					might fall asleep while examining a patient:
				Time points NR	13.2 vs. 7.7, p=0.001; look forward to sleeping
				•	at grand rounds: 10.4 vs. 7.4, p=0.002;

Study	Study	Exposures or intervention		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	
		time points		time points	
					3) No difference in ESS score for those who
					agree/strongly agree vs. other response: work
					is unaffected by sleep loss and fatigue, thinking
					is unaffected by sleep loss, sleep loss and
					fatigue affect my medical decisions, have
					heard of others making medical errors due to
					sleep loss and fatigue, never make errors in
					prescriptions on post-call days, have made
					medical errors because of sleep loss and
					fatigue;
					4) Higher ESS scores predicted by impact score
					in multivariate regression including personal
					and work-related factors: $\beta$ =0.11, p=0.005.
Heponiemi, 2014	CS	Sleeping problems: 4-item	Mean±SD (range) score:	Work ability: Work Ability	1) On-call duty had an indirect effect on work
		Jenkins Scale on 6-point	2.30±1.00 (1-6)	Index on scale from 1 (could	ability (R <sup>2</sup> =0.11, 95% CI: -0.122, -0.031,
RoB: low		scale from 1 (never) to 6		not work at all) to 10 (best	p<0.001) through two mediators (work
		(every night)		work ability)	interference with family, sleeping problems);
					2) Sleeping problems inversely associated with
		Assessed in 2006		Assessed in 2010	work ability, $\beta$ =-0.29, p<0.001.
Kanieta, 2011	CS	Sleep hours: self-reported	Insufficient rest: 32.5%;	Self-reported medical	1) Prevalence of medical incidents (% (95%
		(continuous)	Daytime sleepiness: 3.5%;	incidents: 4-point scale from	CI)): sleep deprived (26.8% (24.2, 29.4)) vs. no
RoB: unclear		Sleepiness and sleep	Insomnia: 20.0%;	1 (never) to 4 (often)	(15.2% (13.7, 16.7)), p<0.01; insomnia (24.8%
		difficulties: 5-point scale	Sleep time (mean±SD min):		(21.6, 28.0)) insomnia vs. not (17.6% (16.2,
		from 1 (never) to 5 (always);	279.8±60.9	Assessed for the past month	19.0)), p<0.01; ≥6h sleep (18.3% (16.8, 19.8))
		Insomnia: ≥3 sleep			vs. <6h (21.7% (18.8, 24.6)), p=0.03;
		difficulties			2) Predictors of medical incidents in
					multivariate model including personal and
		Assessed for the past month			work-related factors (OR (95% CI)): lacking res
					due to sleep deprivation vs. not (1.65 (1.33-
					2.04)), p<0.01); insomnia vs. not (1.45 (1.16-
					1.82), p<0.01); ns for sleep hours.

Study	Study	Exposures or intervention		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	_
		time points		time points	
Sexton, 2001	CS	Fatigue as a factor	NR	Performance effectiveness	1) "When fatigued, I perform effectively during
		impacting performance		measured by 1 question:	critical phases of operations/patient care":
RoB: high				agree, neutral, disagree	Anesthetic: 47% agree; 15% neutral; 38%
		Time points NR			disagree;
				Time points NR	Surgical: 70% agree; 12% neutral; 18%
					disagree.
Shirom, 2006	CS	Tiredness and exhaustion:	NR	Quality of care: Adapted 15-	1) Quality of care positively predicted by
		SMBM Physician Fatigue		item SERVQUAL with a 5-	fatigue in a model incorporating several other
RoB: low		Subscale on a 7-point scale		point Likert scale from 1	components of burnout, $\beta$ =0.17, p<0.05.
		from 1 (almost never) to 7		(very small extent) to 5 (very	
		(always)		large extent)	
		Time points NR		Time points NR	
Smith, 2017	CS	Sleep deprivation: self-	NR	Perceived competence: self-	Some physicians indicated that continual
		reported via open-ended		reported via open-ended	tiredness and exhaustion led to concerns that
RoB: moderate		comments		comments	it would affect their competence; some felt
					that professional performance was
		Time points NR		Time points NR	compromised at times of physical and mental
					fatigue.
Tanti, 2017	CS	Fatique: questionnaire on	NR	Prescribing errors:	Perception of the contribution of fatigue to
		contributors to prescribing		questionnaire on	prescribing errors differed by physician type
RoB: high		errors, with a 5-point Likert		contributors to prescribing	(p<0.05): 34% of community doctors, 96%
		scale (very high to very low		errors, with a 5-point Likert	hospital doctors, 8% of office-working doctors
		association)		scale (very high to very low	perceived a very high or high association
				association)	between fatigue and prescribing errors.
		Time points NR			
				Time points NR	

<sup>&</sup>lt;sup>a</sup>Includes studies of anesthetists, where these were physicians.

<sup>&</sup>lt;sup>b</sup>Includes primary care physicians, internal medicine physicians, and general practitioners.

ACC: aortic cross-clamp time; BA: before-after; CABG: cardiopulmonary bypass time; CCPT II: Connor's Continuous Performance Test II; CI: confidence interval; CO: cohort; CS: cross-sectional; ER: emergency; ESS: Epworth Sleepiness Scale; h: hour(s); HVLT: Hopkin's Verbal Learning Task; Hz: Hertz; ICU: intensive care unit; IQR: interquartile range; KSS: Karolinska Sleep Scale; min: minutes; ms: millisecond(s); N-back: Dual N-back test; NA: not applicable; NR: not reported; ns: not statistically significant; OR: odds ratio; PSQI: Pittsburgh Sleep Quality Index; PVT: Psychomotor vigilance Performance Task; RR: risk ratio; RCT: randomized controlled trial; RoB: Risk of Bias; SD: standard deviation; SE: standard error; SERVQUAL: Service Quality Measure; SMBM: Shirom-Melamed Burnout Measure; TP: Toulouse-Piéron test; TS: time series; US: United States of America; vs.: versus

Patient outcomes related to fatigue or sleep restriction among physicians in independent practice

Study	Study	Exposures		Outcome Measures	Associations between exposure and outcome
Risk of Bias (RoB)	design	Intervention or assessment	Baseline	Assessment scale and time	_
		scale and time points		points	
Surgeons					
Chu, 2011	СО	Sleep deprivation: moderate	Of 4,047 procedures, 83	Chart review: mortality, surgical	1) 0-3 vs. 3-6 vs. >6 hours of sleep: No
		(3-6 h) or severe (<3-h) sleep	(2.1%) performed by	complications, length of stay	difference in incidence of mortality, incidence
RoB: low		deprivation the night before	severely sleep-deprived,		of 10 major complications (except septicemia,
		surgery (self-reported hours)	1,595 (39.4%) by	Assessed during and post-	3.6% vs. 0.9% vs. 0.8%, p=0.03), ICU length of
			moderately sleep-	surgery	stay; in-hospital length of stay (days): 7.0 vs.
			deprived surgeons		6.0 vs. 7.0, p<0.001.
Ellman, 2004	СО	Sleep deprivation: performed	Of 6,751 procedures,	Chart review: mortality, surgical	1) Sleep deprived vs. non-sleep deprived: no
		a case starting 22:00 to 05:00,	339 (5%) were	complications, length of stay	difference in mortality, need for blood
RoB: low		or ending 22:00 to 07:30 and	performed by sleep		products, complications (operative,
		performed a subsequent case	deprived surgeons	Assessed during and post-	neurologic, renal, infectious, pulmonary), in-
		in the next 24-h		surgery	hospital length of stay.
Govindarajan,	СО	Sleep deprivation: treated	NR	Chart review: mortality, surgical	1) Sleep deprived vs. non-sleep deprived: no
2015		patients from midnight to		complications, readmission,	difference in mortality, surgical complications,
		07:00 and performed a		length of stay	readmissions within 30 days, or length of stay.
RoB: low		subsequent case on the same			
		day		Assessed during and post-	
				surgery	
Rothschild, 2009	СО	Sleep deprivation: daytime	NR	Chart review: frequency of	1) Post-nighttime vs. control: no difference in
		procedures following an		adverse surgical complications	number of procedures with complications,
RoB: low		overnight procedure;			total number of complications, preventable
		Sleep opportunity: 0-6h, <6h		Assessed during and post-	complications, type of complications;
				surgery	2) Operating room procedures with
					complications, OR (95% CI): 8.5% for 0-6h
					sleep vs. 3.1% for >6h sleep, 2.70 (1.13-6.48),
					p=0.03;
					3) All procedures with complications, OR (95%
					CI): 6.2% for 0-6h sleep vs. 3.4% for >6h sleep,
					1.72 (1.02-2.89), p=0.04.

Study	Study	Exposures		Outcome Measures	Associations between exposure and outcome	
Risk of Bias (RoB)	design	Intervention or assessment	Baseline	Assessment scale and time	_	
		scale and time points		points		
Schieman, 2007	СО	Fatigue: surgeon billed for	Of 270 procedures, 22	Chart review: surgical	1) Fatigued vs. non-fatigued surgeons: no	
		clinical work after 22:00 the	(8%) were performed by	complications, length of stay,	difference in intra- or post-operative	
RoB: low		night before surgery	fatigued surgeons	mortality, cancer recurrence	complication rate, length of stay, in-hospital	
					length of stay, cancer recurrence.	
				Assessed during and post-		
		A TOTAL CONTRACTOR OF THE PROPERTY OF THE PROP		surgery		
Vinden, 2014	CO	Sleep deprivation (at risk):	Of 94,183 surgeries,	Chart review: conversion to	1) At risk vs. not at risk surgeon: no difference	
		surgeon worked 00:00 to	2,078 (2.2%) were	open procedure (from	in incidence of conversion to open procedure,	
RoB: low		07:00 and performed surgery	performed by surgeons	laparoscopic), iatrogenic	iatrogenic injuries, mortality, in either	
		07:00 to 18:00	who were 'at risk'	injuries, mortality	univariate or multivariate analyses.	
				Assessed during and post-		
				surgery		
Obstetricians						
Rothschild, 2009	СО	Sleep deprivation: daytime	NR	Chart review: frequency of	1) Post-nighttime vs. control: no difference in	
		procedures following an		adverse obstetric complications	number of procedures with complications,	
RoB: low		overnight procedure;			total complications, preventable	
		Sleep opportunity: 0-6h, <6h		Assessed during and post-	complications, type of complications;	
				delivery	2) No association between sleep deprivation	
					and proportion of procedures with	
					complications, nor difference for 0-6h vs. >6h	
					of sleep opportunity.	

CI: confidence interval; CO: cohort; h: hours; NR: not reported; OR: odds ratio; RoB: Risk of Bias; SD: standard deviation; US: United States of America; vs.: versus

## **Dichotomous outcomes**

Outcome or subgroup	Number	Number of	Pooled risk ratio	l <sup>2</sup>
	of studies	participants	(95% CI)	
1.1 Patient mortality	5	60,436	0.98 (0.84, 1.15)	0%
1.2 Intra-operative complications	3	19,798	1.35 (0.82, 2.21)	82%
1.2.1 Surgical procedure	3 <sup>a</sup>	14,896	1.37 (0.65, 2.87)	88%
1.2.2 Obstetric procedure	<b>1</b> <sup>a</sup>	4,902	1.21 (0.83, 1.78)	NA
1.3 Post-operative complications	5	60,201	0.99 (0.95, 1.03)	0%

<sup>&</sup>lt;sup>a</sup> Rothschild, 2009 is represented in both analyses

# **Continuous outcomes**

Outcome or subgroup	Number	Number of	Pooled mean	l <sup>2</sup>
	of studies	participants	difference (95% CI)	
1.4 Operating time (minutes)	4	50,046	-0.14 (-1.60, 1.33)	0%
1.5 Length of hospital stay (days)	4	50,046	-0.33 (-1.03, 0.36)	86%
1.5.1 Cardiac surgeries	2	10,798	-0.43 (-1.55, 0.69)	84%
1.5.2 Elective surgeries	1	38,978	0.00 (-0.07, 0.07)	NA
1.5.3 Anterior resection for anal cancer	1	270	-2.10 (-5.98, 1.78)	NA

CPBT: cardiopulmonary bypass time; NA: not applicable

# 1.1 Patient mortality

	Sleep de	prived	Not sleep d	eprived		Risk Ratio			Risk	Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI	Year		M-H, Rand	om, 95% CI		
Ellman 2004	17	339	296	6412	10.4%	1.09 [0.67, 1.75]	2004		_	-		
Schieman 2008	0	22	3	248	0.3%	1.55 [0.08, 29.03]	2008			-		
Chu 2011	47	1678	80	2369	18.8%	0.83 [0.58, 1.18]	2011			<del> </del>		
Vinden 2014	2	2078	7	8312	1.0%	1.14 [0.24, 5.50]	2014			<del>-</del>		
Govindarajan 2015	224	19489	222	19489	69.5%	1.01 [0.84, 1.21]	2015		•			
Total (95% CI)		23606		36830	100.0%	0.98 [0.84, 1.15]			•			
Total events	290		608									
Heterogeneity: Tau <sup>2</sup> =	0.00; Chi <sup>2</sup>	= 1.25, d	f = 4 (P = 0.8)	7); I <sup>2</sup> = 0%				0.04	01	<del>                                     </del>		400
Test for overall effect:	Z = 0.23 (F	P = 0.82)						0.01 Fav	vors sleep deprived	Favors non-slee	) ep deprive	100 ed

# Sensitivity analysis using highest possible number of events for Vinden 2014

	Sleep de	prived	Not sleep d	eprived		Risk Ratio		Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI	Year	M-H, Random, 95% CI
Ellman 2004	17	339	296	6412	13.9%	1.09 [0.67, 1.75]	2004	<del>-</del>
Schieman 2008	0	22	3	248	0.4%	1.55 [0.08, 29.03]	2008	<del></del>
Chu 2011	47	1678	80	2369	23.3%	0.83 [0.58, 1.18]	2011	<del></del>
Vinden 2014	5	2078	7	8312	2.6%	2.86 [0.91, 8.99]	2014	<del></del>
Govindarajan 2015	224	19489	222	19489	59.8%	1.01 [0.84, 1.21]	2015	•
Total (95% CI)		23606		36830	100.0%	1.00 [0.83, 1.21]		<b>+</b>
Total events	293		608					
Heterogeneity: Tau <sup>2</sup> =	0.01; Chi²	= 4.50, d	f = 4 (P = 0.34)	4); I <sup>2</sup> = 119	%			0.01 0.1 10 100
Test for overall effect:	Z = 0.03 (F	P = 0.98)						Favors sleep deprived Favors non-sleep deprived

# Sensitivity analysis using lowest possible number of events for Vinden 2014



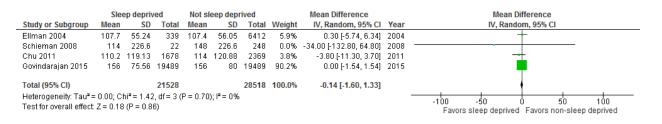
# 1.2 Intra-operative complications

-
<u> </u>
5
prived

# 1.3 Post-operative complications

	Sleep de	prived	Not sleep d	eprived		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI	M-H, Random, 95% CI
Chu 2011	259	1678	404	2369	7.8%	0.91 [0.78, 1.04]	-+-
Ellman 2004	29	339	480	6412	1.2%	1.14 [0.80, 1.64]	<del></del>
Govindarajan 2015	3527	19489	3543	19489	89.4%	1.00 [0.95, 1.04]	
Schieman 2008	12	22	164	248	1.0%	0.82 [0.56, 1.22]	<del></del>
Vinden 2014	14	2031	72	8124	0.5%	0.78 [0.44, 1.38]	
Total (95% CI)		23559		36642	100.0%	0.99 [0.95, 1.03]	•
Total events	3841		4663				
Heterogeneity: Tau <sup>2</sup> =	= 0.00; Chi²	' = 3.69, d	f = 4 (P = 0.49)	5); I² = 0%			
Test for overall effect							0.2 0.5 1 2 5 Favors sleep deprived Favors non-sleep deprived

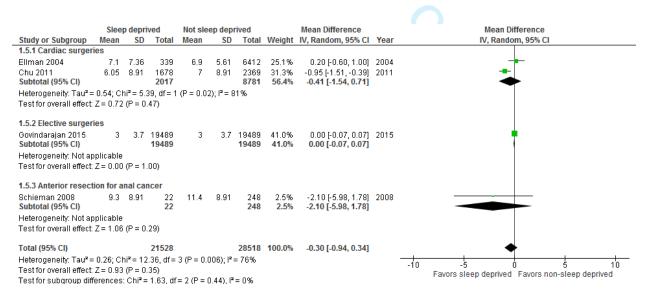
Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies



## 1.5 Length of hospital stay (days)

Sleep   Slee											
1.5.1 Cardiac surgeries  Ellman 2004 7.1 7.36 339 6.9 5.61 6412 25.4% 0.20 [-0.60, 1.00] 2004  Chu 2011 6.05 6.66 1678 7 6.07 2369 33.8% -0.95 [-1.35, -0.55] 2011  Subtotal (95% CI) 2017 88781 59.2% -0.43 [-1.55, 0.69]  Heterogeneity: Tau* = 0.56; Chi* = 6.40, df=1 (P = 0.01); P = 84%  Test for overall effect: Z = 0.75 (P = 0.45)  1.5.2 Elective surgeries  Govindarajan 2015 3 3.7 19489 3 3.7 19489 37.9% 0.00 [-0.07, 0.07] 2015  Subtotal (95% CI) 19489 19489 37.9% 0.00 [-0.07, 0.07]  Heterogeneity: Not applicable  Test for overall effect: Z = 0.00 (P = 1.00)  1.5.3 Anterior resection for anal cancer  Schieman 2008 9.3 8.91 22 11.4 8.91 248 2.9% -2.10 [-5.98, 1.78] 2008  Subtotal (95% CI) 22 248 2.9% -2.10 [-5.98, 1.78]  Heterogeneity: Not applicable  Test for overall effect: Z = 1.06 (P = 0.29)  Total (95% CI) 21528 28518 100.0% -0.33 [-1.03, 0.36]  Heterogeneity: Tau* = 0.33; Chi* = 22.21, df = 3 (P < 0.0001); I* = 86%  Test for overall effect: Z = 0.94 (P = 0.35)  Test for overall effect: Z = 0.94 (P = 0.36) Favors sleep deprived Favors non-sleep deprived		Slee	p depr	ived	Not sle	ep depi	rived		Mean Difference		Mean Difference
Ellman 2004 7.1 7.36 339 6.9 5.61 6412 25.4% 0.20 [-0.60, 1.00] 2004 Chu 2011 6.05 6.66 1678 7 6.07 2369 33.8% -0.95 [-1.35, -0.55] 2011 Subtotal (95% CI)	Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	Year	IV, Random, 95% CI
Chu 2011 6.05 6.66 1678 7 6.07 2369 33.8% -0.95 [-1.35, -0.55] 2011  Subtotal (95% CI) 2017 8781 59.2% -0.43 [-1.55, 0.69]  Heterogeneity. Tau² = 0.56; Chi² = 6.40, df = 1 (P = 0.01); i² = 84%  Test for overall effect Z = 0.75 (P = 0.45)  1.5.2 Elective surgeries  Govindarajan 2015 3 3.7 19489 37.9% 0.00 [-0.07, 0.07] 2015  Subtotal (95% CI) 19489 19489 37.9% 0.00 [-0.07, 0.07]  Heterogeneity. Not applicable  Test for overall effect Z = 0.00 (P = 1.00)  1.5.3 Anterior resection for anal cancer  Schieman 2008 9.3 8.91 22 11.4 8.91 248 2.9% -2.10 [-5.98, 1.78] 2008  Subtotal (95% CI) 22 248 2.9% -2.10 [-5.98, 1.78]  Heterogeneity. Not applicable  Test for overall effect Z = 0.33; Chi² = 22.21, df = 3 (P < 0.0001); i² = 86%  Test for overall effect Z = 0.94 (P = 0.35)  Test for subgroup differences: Chi² = 1.68 df = 2 (P = 0.43), i² = 0%  Favors sleep deprived  Favors non-sleep deprived	1.5.1 Cardiac surger	ies									
Subtotal (95% CI) 2017 8781 59.2% -0.43 [-1.55, 0.69]  Heterogeneity: Tau* = 0.56; Chi* = 6.40, df = 1 (P = 0.01); i* = 84%  Test for overall effect: Z = 0.75 (P = 0.45)  1.5.2 Elective surgeries  Govindarajan 2015 3 3.7 19489 3 7.9% 0.00 [-0.07, 0.07] 2015  Subtotal (95% CI) 19489 19489 37.9% 0.00 [-0.07, 0.07]  Heterogeneity: Not applicable  Test for overall effect: Z = 0.00 (P = 1.00)  1.5.3 Anterior resection for anal cancer  Schieman 2008 9.3 8.91 22 11.4 8.91 248 2.9% -2.10 [-5.98, 1.78] 2008  Subtotal (95% CI) 22 248 2.9% -2.10 [-5.98, 1.78]  Heterogeneity: Not applicable  Test for overall effect: Z = 1.06 (P = 0.29)  Total (95% CI) 21528 28518 100.0% -0.33 [-1.03, 0.36]  Heterogeneity: Tau* = 0.33; Chi* = 22.21, df = 3 (P < 0.0001); i* = 86%  Test for overall effect: Z = 0.94 (P = 0.35)  Test for subgroup differences: Chi* = 1.68, df = 2 (P = 0.43) i* = 0.%  Favors sleep deprived  Favors non-sleep deprived	Ellman 2004	7.1	7.36	339	6.9	5.61	6412	25.4%	0.20 [-0.60, 1.00]	2004	+
Heterogeneity: Tau² = 0.56; Chi² = 6.40, df = 1 (P = 0.01); i² = 84%  Test for overall effect: Z = 0.75 (P = 0.45)  1.5.2 Elective surgeries  Govindarajan 2015	Chu 2011	6.05	6.66	1678	7	6.07	2369	33.8%	-0.95 [-1.35, -0.55]	2011	<b>+</b> _
Test for overall effect: Z = 0.75 (P = 0.45)  1.5.2 Elective surgeries  Govindarajan 2015	Subtotal (95% CI)			2017			8781	59.2%	-0.43 [-1.55, 0.69]		•
1.5.2 Elective surgeries Govindarajan 2015	Heterogeneity: Tau² =	0.56; Cl	hi² = 6.	40, df = 1	(P = 0.0)	1); $I^2 = 8$	14%				
Govindarajan 2015 3 3.7 19489 3 3.7 19489 37.9% 0.00 [-0.07, 0.07] 2015 Subtotal (95% CI) 19489 19489 37.9% 0.00 [-0.07, 0.07] Heterogeneity. Not applicable Test for overall effect: Z = 0.00 (P = 1.00)  1.5.3 Anterior resection for anal cancer Schieman 2008 9.3 8.91 22 11.4 8.91 248 2.9% -2.10 [-5.98, 1.78] 2008 Subtotal (95% CI) 22 248 2.9% -2.10 [-5.98, 1.78] Heterogeneity. Not applicable Test for overall effect: Z = 1.06 (P = 0.29)  Total (95% CI) 21528 28518 100.0% -0.33 [-1.03, 0.36] Heterogeneity. Tau² = 0.33; Chi² = 22.21, df = 3 (P < 0.0001); i² = 86% Test for overall effect: Z = 0.94 (P = 0.35) Test for subgroup differences: Chi² = 1.68, df = 2 (P = 0.43) i² = 0%  Favors sleep deprived	Test for overall effect:	Z = 0.75	(P = 0	1.45)							
Subtotal (95% CI) 19489 19489 37.9% 0.00 [-0.07, 0.07]  Heterogeneity: Not applicable Test for overall effect: Z = 0.00 (P = 1.00)  1.5.3 Anterior resection for anal cancer Schieman 2008 9.3 8.91 22 11.4 8.91 248 2.9% -2.10 [-5.98, 1.78] 2008 Subtotal (95% CI) 22 248 2.9% -2.10 [-5.98, 1.78]  Heterogeneity: Not applicable Test for overall effect: Z = 1.06 (P = 0.29)  Total (95% CI) 21528 28518 100.0% -0.33 [-1.03, 0.36]  Heterogeneity: Tau² = 0.33; Chi² = 22.21, df = 3 (P < 0.0001); i² = 86% Test for overall effect: Z = 0.94 (P = 0.35)  Test for overall effect: Z = 0.94 (P = 0.35)  Test for subgroup differences: Chi² = 1.68, df = 2 (P = 0.43) i² = 0%	1.5.2 Elective surger	ies									
Heterogeneity: Not applicable Test for overall effect: Z = 0.00 (P = 1.00)  1.5.3 Anterior resection for anal cancer Schieman 2008 9.3 8.91 22 11.4 8.91 248 2.9% -2.10 [-5.98, 1.78] 2008 Subtotal (95% CI) 22 248 2.9% -2.10 [-5.98, 1.78] Heterogeneity: Not applicable Test for overall effect: Z = 1.06 (P = 0.29)  Total (95% CI) 21528 28518 100.0% -0.33 [-1.03, 0.36] Heterogeneity: Tau² = 0.33; Chi² = 22.21, df = 3 (P < 0.0001); i² = 86% Test for overall effect: Z = 0.94 (P = 0.35) Test for subgroup differences: Chi² = 1.88, df = 2 (P = 0.43) i² = 0%  Favors sleep deprived	Govindarajan 2015	3	3.7	19489	3	3.7	19489	37.9%	0.00 [-0.07, 0.07]	2015	•
Test for overall effect: Z = 0.00 (P = 1.00)  1.5.3 Anterior resection for anal cancer Schieman 2008 9.3 8.91 22 11.4 8.91 248 2.9% -2.10 [-5.98, 1.78] 2008 Subtotal (95% CI) 22 248 2.9% -2.10 [-5.98, 1.78] Heterogeneity: Not applicable Test for overall effect: Z = 1.06 (P = 0.29)  Total (95% CI) 21528 28518 100.0% -0.33 [-1.03, 0.36] Heterogeneity: Tau² = 0.33; Chi² = 22.21, df = 3 (P < 0.0001); i² = 86% Test for overall effect: Z = 0.94 (P = 0.35) Test for subgroup differences: Chi² = 1.68, df = 2 (P = 0.43) i² = 0%  Favors sleep deprived Favors non-sleep deprived	Subtotal (95% CI)			19489			19489	37.9%	0.00 [-0.07, 0.07]		
1.5.3 Anterior resection for anal cancer  Schieman 2008 9.3 8.91 22 11.4 8.91 248 2.9% -2.10 [-5.98, 1.78] 2008  Subtotal (95% CI) 22 248 2.9% -2.10 [-5.98, 1.78]  Heterogeneity: Not applicable  Test for overall effect: Z = 1.06 (P = 0.29)  Total (95% CI) 21528 28518 100.0% -0.33 [-1.03, 0.36]  Heterogeneity: Tau² = 0.33; Chi² = 22.21, df = 3 (P < 0.0001); i² = 86%  Test for overall effect: Z = 0.94 (P = 0.35)  Test for subgroup differences: Chi² = 1.88, df = 2 (P = 0.43) i² = 0%  Favors sleep deprived	Heterogeneity: Not as	oplicable									
Schieman 2008 9.3 8.91 22 11.4 8.91 248 2.9% -2.10 [-5.98, 1.78] 2008  Subtotal (95% CI) 22 248 2.9% -2.10 [-5.98, 1.78] 2008  Heterogeneity: Not applicable Test for overall effect: Z = 1.06 (P = 0.29)  Total (95% CI) 21528 28518 100.0% -0.33 [-1.03, 0.36]  Heterogeneity: Tau² = 0.33; Chi² = 22.21, df = 3 (P < 0.0001); i² = 86%  Test for overall effect: Z = 0.94 (P = 0.35)  Test for subgroup differences: Chi² = 1.68, df = 2 (P = 0.43) i² = 0%  Favors sleep deprived	Test for overall effect:	Z = 0.00	(P = 1	.00)							
Subtotal (95% CI) 22 248 2.9% -2.10 [-5.98, 1.78]  Heterogeneity: Not applicable Test for overall effect: Z = 1.06 (P = 0.29)  Total (95% CI) 21528 28518 100.0% -0.33 [-1.03, 0.36]  Heterogeneity: Tau² = 0.33; Chi² = 22.21, df = 3 (P < 0.0001); i² = 86%  Test for overall effect: Z = 0.94 (P = 0.35)  Test for subgroup differences: Chi² = 1.88, df = 2 (P = 0.43) i² = 0%  Favors sleep deprived	1.5.3 Anterior resect	tion for a	nal ca	ncer							
Heterogeneity: Not applicable Test for overall effect: Z = 1.06 (P = 0.29)  Total (95% CI)  Heterogeneity: Tau² = 0.33; Chi² = 22.21, df = 3 (P < 0.0001); i² = 86% Test for overall effect: Z = 0.94 (P = 0.35) Test for subgroup differences: Chi² = 1.88, df = 2 (P = 0.43) i² = 0%  Test for subgroup differences: Chi² = 1.88, df = 2 (P = 0.43) i² = 0%	Schieman 2008	9.3	8.91	22	11.4	8.91	248	2.9%	-2.10 [-5.98, 1.78]	2008	<del></del>
Test for overall effect: Z = 1.06 (P = 0.29)  Total (95% CI)  100.08  21528  28518  28518  100.08  -0.33 [-1.03, 0.36]  Heterogeneity: Tau² = 0.33; Chi² = 22.21, df = 3 (P < 0.0001); i² = 86%  Test for overall effect: Z = 0.94 (P = 0.35)  Test for subgroup differences: Chi² = 1.88. df = 2 (P = 0.43) i² = 0%  Favors sleep deprived  Favors non-sleep deprived	Subtotal (95% CI)			22			248	2.9%	-2.10 [-5.98, 1.78]		
Total (95% CI) 21528 28518 100.0% -0.33 [-1.03, 0.36]  Heterogeneity: Tau² = 0.33; Chi² = 22.21, df = 3 (P < 0.0001); i² = 86% 10	Heterogeneity: Not as	oplicable									
Heterogeneity: Tau² = 0.33; Chi² = 22.21, df = 3 (P < 0.0001); i² = 86%  Test for overall effect: Z = 0.94 (P = 0.35)  Test for subgroup differences: Chi² = 1.68, df = 2 (P = 0.43), i² = 0%  Favors sleep deprived  Favors non-sleep deprived	Test for overall effect	Z=1.06	(P = 0	.29)							
Test for overall effect: Z = 0.94 (P = 0.35)  Test for subgroup differences: Chi² = 1.68, df = 2 (P = 0.43), l² = 0%	Total (95% CI)			21528			28518	100.0%	-0.33 [-1.03, 0.36]		•
Test for overall effect: Z = 0.94 (P = 0.35)  Test for subgroup differences: Chi² = 1.68, df = 2 (P = 0.43), l² = 0%	Heterogeneity: Tau <sup>2</sup> =	= 0.33; CI	hi² = 22	2.21. df=	3 (P < 0.0	0001): P	²= 86%			-	<del></del>
Test for subgroup differences: Chi² = 1.68. df = 2 (P = 0.43), I² = 0%	- '										
					= 2 (P = I	0.43), I²	= 0%				Favors sleep deprived Favors non-sleep deprived

## Sensitivity analysis imputing the highest standard deviation



# Sensitivity analysis imputing the lowest standard deviation

State on Salaman	Sleep deprived	Not sleep deprived		Mean Difference	Mean Difference
Study or Subgroup  1.5.1 Cardiac surger	Mean SD Total	Mean SD To	itai vveigni	t IV, Random, 95% CI	I Year IV, Random, 95% CI
Ellman 2004 Chu 2011	7.1 7.36 339 6.05 3.7 1678	7 3.7 23	12 25.7% 69 34.9%	-0.95 [-1.18, -0.72]	2011
Subtotal (95% CI)  Heterogeneity: Tau <sup>2</sup> =  Test for overall effect:	2017 0.57; Chi² = 7.40, df = Z = 0.77 (P = 0.44)		781 60.6%	-0.44 [-1.56, 0.68]	
1.5.2 Elective surger	ios.				
Govindarajan 2015 Subtotal (95% CI)	3 3.7 19489 <b>1948</b> 9				
Heterogeneity: Not ap Test for overall effect:					
1.5.3 Anterior resect	ion for anal cancer				
Schieman 2008 Subtotal (95% CI)	9.3 8.91 22 22		48 3.4% 48 <b>3.4</b> %		
Heterogeneity: Not ap Test for overall effect:					
Total (95% CI) Heterogeneity: Tau² =	21528 0.40; Chi² = 60.35, df		35%	-0.35 [-1.10, 0.40]	1
Test for overall effect: Test for subgroup diff	Z= 0.92 (P = 0.36) erences: Chi² = 1.71, c	lf = 2 (P = 0.42), I <sup>2</sup> = 09	6		Favors sleep deprived Favors non-sleep deprived



age 87 of 89		BMJ Open BMJ open				
Appendix 1. PRISMA checklist  Open - 2018						
Section/topic	#	Checklist item the control of the co	Reported on page #			
TITLE		ding 2				
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1			
ABSTRACT		use:				
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; state gibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; concessions and implications of key findings; systematic review registration number.	2			
INTRODUCTION		Do tex				
Rationale	3	Describe the rationale for the review in the context of what is already known.	4-5			
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5			
METHODS		m BES)				
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if a like, provide registration information including registration number.	5			
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years encourage, publication status) used as criteria for eligibility, giving rationale.	6			
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors identify additional studies) in the search and date last searched.	5-6			
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that income a search strategy for at least one database, including any limits used, such that income a search strategy for at least one database, including any limits used, such that is could be repeated.	Supplementary file 1			
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and included in the meta-analysis).	6-7			
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.	7			
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assum ions and simplifications made.	7			
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.	7			
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	8			
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures for each meta-analysis.	8			
4 5 6		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml				



45 46 47

		BMJ Open  BMJ open	Page 88 of
Appendix 1. PRISM	MA chec	BMJ Open  BMJ Open  cted by copyrig	
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 is a specified by the complete specified by the co	8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression killing methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression killing methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression killing methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression killing methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression killing methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression killing methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression killing methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression killing methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression killing methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression killing methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression killing methods of additional analyses (e.g., sensitivity or subgroup analyses).	8
RESULTS		relation of the state of the st	
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reason exclusions at each stage, ideally with a flow diagram.	8, Figure 1
5 Study characteristics 6 7	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follows for period) and provide the citations.	8-11, Table 1, Supplementary file 2
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see it ).	11, Supplementary file 3
Results of individual studies Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary date for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	p. 12-18; Supplementary file 4; figures 2-6
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consecutive.	p. 12-18, figures 2-6
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Not applicable
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Supplementary file 5
DISCUSSION		9, 2 blog	
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider the relevance to key groups (e.g., healthcare providers, users, and policy makers).	18-19
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplet research, reporting bias).	19-20
9 Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for turn research.	20
FUNDING		og r	
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role objective systematic review.	21
<del></del>		For near review only, http://hmienen.hmi.com/cite/ahout/guidelines.yhtml	•

\*\*September 2018. Downstanded First Interpretability. Attending and similar technologies.

\*\*September 2019. Downstanded Titrum http://bm/popen.bm/scom/ on June 8, 2003 at Agency (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses. Including for users information, visit: <a href="http://www.prisma-statement.org">www.prisma-statement.org</a>. Page 2 of 2

\*\*September 2018. Downstanded First Interpretability.//bm/popen.bm/scom/ on June 8, 2003 at Agency (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses. Including for users and data mining. At training, and similar technologies. 

# **BMJ Open**

# The impact of fatigue and insufficient sleep on physician and patient outcomes: A systematic review

Journal:	BMJ Open				
Manuscript ID	bmjopen-2018-021967.R2				
Article Type:	Research				
Date Submitted by the Author:	28-Jun-2018				
Complete List of Authors:	Gates, Michelle; University of Alberta, Pediatrics Wingert, Aireen; University of Alberta, Pediatrics Featherstone, Robin; University of Alberta, Pediatrics Samuels, Charles; Centre for Sleep and Human Performance Simon, Christopher; Canadian Medical Association Dyson, Michele; University of Alberta, Pediatrics				
<b>Primary Subject Heading</b> :	Occupational and environmental medicine				
Secondary Subject Heading:	Evidence based practice, Health services research, Medical education and training, Occupational and environmental medicine				
Keywords:	Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, SLEEP MEDICINE				

SCHOLARONE™ Manuscripts



mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

The impact of fatigue and insufficient sleep on physician and patient outcomes: A systematic review

Michelle Gates, Aireen Wingert, Robin Featherstone, Charles Samuels, Christopher Simon, Michele P Dyson

**Michelle Gates**, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-486C, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Project coordinator

**Aireen Wingert**, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-490B, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Project coordinator

**Robin Featherstone**, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-486D, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Research librarian

**Charles Samuels**, Centre for Sleep and Human Performance, Suite 106, 51 Sunpark Drive SE, Calgary, Alberta, Canada, T2X 3V4
Physician and medical director

**Christopher Simon**, Canadian Medical Association, 1867 Alta Vista Drive, Ottawa, Ontario, Canada, K1G 5W8

Advisor, physician health

Michele P Dyson, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-474, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Assistant professor

Correspondence to: MP Dyson mdyson@ualberta.ca

Word count (main text): 5,929

**Objectives:** For physicians in independent practice, we synthesized evidence on the (a) impacts of insufficient sleep and fatigue on health and performance, and patient safety; (b) effectiveness of interventions targeting insufficient sleep and fatigue.

**Design:** We systematically reviewed online literature. After piloting, one reviewer selected studies by title and abstract; full texts were then reviewed in duplicate. One reviewer extracted data; another verified a random 10% sample. Two reviewers assessed risk of bias. We pooled findings via meta-analysis when appropriate, or narratively.

**Data sources:** We searched Medline, Embase, PsycINFO, CINAHL and PubMed for published studies in April 2016; Medline was updated in November 2017. We searched Embase for conference proceedings, and hand-searched meeting abstracts, association and foundation websites.

**Eligibility criteria for selecting studies:** English or French language primary research studies published from 2000-2017 examining the effect of fatigue or sleep-related exposures or interventions on any outcome among physicians in independent practice and their patients.

**Results:** Of 16,154 records identified, we included 47 quantitative studies of variable quality. 28 studies showed associations between fatigue or insufficient sleep and physician health and well-being outcomes. 21 studies showed no association with surgical performance, and mixed findings for psychomotor performance, work performance, and medical errors. We pooled data from six cohort studies for patient outcomes. For sleep deprived versus non-sleep deprived surgeons, we found no difference in patient mortality (n = 60,436, RR 0.98, 95% CI 0.84 to 1.15, p = 0.82,  $I^2 = 0\%$ ) nor postoperative complications (n = 60,201, RR 0.99, 95% CI 0.95 to 1.03,  $I^2 = 0\%$ ). The findings for intraoperative complications and length of stay were considerably heterogeneous.

**Conclusions:** Fatigue and insufficient sleep may be associated with negative physician health outcomes. Current evidence is inadequate to inform practice recommendations.

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text and

- The review was informed by the methods outlined by Cochrane and is reported according to the
   Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.
- The review was limited by the quality of the included studies, which was often poor. We could not draw definitive conclusions due to methodological weaknesses and heterogeneous outcome measures in the included studies.
- We have focused on evidence from high income countries; our findings may not be generalizable to other settings.

The working hours of physicians have been a topic of debate for many years.[1] Beginning in the late 1980s, evidence indicating that medical resident fatigue could negatively impact their cognitive functioning and performance, resulting in an increased risk of medical error, began to accumulate.[2] In response, by the early 2000s physicians' regulatory bodies worldwide began to take action toward restricting the work hours of medical residents and ensuring adequate time for recovery between shifts.[3-5] Since their implementation in the United States by the Accreditation Council for Graduate Medical Education (ACGME), the impact of work hour regulations has been widely researched. Still, evidence for impacts on patient care, resident training and wellbeing remains equivocal.[6-9] This is likely because work hours are only one of many contributors to fatigue and physician wellbeing. In fact, the ACGME has recently reversed the 2011 changes that limited resident work hours to 16 hours per shift and the requirement for 8 hours of time off between shifts. This decision was made in favour of promoting "flexibility" for residency training program work hours and scheduling.

The focus on medical trainees has left physicians in independent practice as a relatively neglected group in research and policy. In Canada, there is no concrete regulation on the hours or patterns in which physicians choose to work.[10] In the absence of clear policies, physicians trained under traditional systems may find it difficult to work shorter hours or take more frequent breaks.[1] Indeed, more than 40% of practicing physicians in the United States work in excess of 80 hours per week.[11] While long work hours remain a cultural norm in medicine, in comparable high-risk industries (e.g., aviation), work patterns and work hours are tightly regulated.[12] The need for similar evidence-based policies in medicine has become a topic of increased interest. Exemplar of this, an evidence-based guideline for fatigue risk management in emergency medical services,[13] informed by a comprehensive set of systematic reviews, has recently been published. For physicians, it has been argued that there is a need to adapt healthcare systems and provide support in identifying the signs of fatigue and mitigating its risks.[1]

Besides potentially affecting patient outcomes, fatigue can impact the health and wellbeing of physicians themselves. Burnout, just one outcome related to fatigue, has been described as epidemic among physicians[14-16] and ultimately affects recruitment and retention of physicians both in community and acute care settings. While the effect of physician wellbeing on the sustainability of healthcare systems has recently received increased attention,[17] evidence-based solutions to burnout

Given this void, we undertook a systematic review focusing broadly on primary research relevant to the Canadian context as a fundamental starting point to examine the effects of fatigue and chronic insufficient sleep on physicians in independent practice, and on interventions to combat these effects. Our review was guided by the following research questions: Among physicians in independent practice, (1) what are the impacts of fatigue and chronic insufficient sleep on physician health, physician performance, and patient safety; and (2) what is the effectiveness of interventions that target fatigue and chronic insufficient sleep, in terms of improving physician and patient outcomes?

#### **METHODS**

#### **Review conduct**

The conduct of this systematic review was guided by Cochrane standards.[22] The research team convened to plan the key research questions and methodology but did not register a formal protocol. The findings are reported in adherence with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) Statement.[23] Ethical approval was not required for this study.

#### **Patient involvement**

Patients were not involved.

## Literature search

An information specialist developed a search strategy that included concepts related to physicians, fatigue and sleep. On 13 April 2016 we searched the following online databases with coverage in the biomedical sciences and psychology: Medline, Embase, PsycINFO, CINAHL and PubMed, limited to English and French language articles published from 2000 to 2016. We updated the Medline search in November 2017, as this database offered the highest precision. Though fatigue among physicians is not a new phenomenon,[2] we limited our search to articles published post-2000 to include studies relevant to current physician practice. Work hour limitations have existed in European countries since 1993, but

disagreements between reviewers were resolved by discussion or third-reviewer consultation when necessary.

#### **Data extraction**

Reviewers used a standardized form to extract data in Microsoft Office Excel (v. 2016, Microsoft Corporation, Redmond, WA). One reviewer independently extracted data from each included study and a second reviewer verified a random 10% sample. Since no major errors or omissions were noted, we did not undertake further verification.

We extracted the following data: country of publication; funding source; study design; inclusion and exclusion criteria; population characteristics (i.e., sample size, age and gender distribution, physician specialty); setting (i.e., physician workplace, urban or rural); exposure or intervention; definition of fatigue or insufficient sleep; sleep and fatigue scales used and timing of measurement; comparators (if applicable); and outcomes.

# Risk of bias appraisal

Two reviewers independently assessed the risk of bias in each included study using standard tools. Disagreements were resolved via discussion or by consulting a third reviewer. We used the Cochrane Risk of Bias tool[22] to assess randomised controlled trials. Adapted versions of the tool developed by the Effective Practice and Organization of Care group[26] were used to assess before-after and time series studies. We used the Newcastle-Ottawa Quality Assessment Scale[27] to appraise cohort studies. We adapted the scale to assess cross-sectional studies and the one non-comparative study.

### **Evidence synthesis**

We considered clinical and methodological heterogeneity in our decision on whether to proceed with meta-analysis for the outcomes identified. For most outcomes, we found high levels of heterogeneity in study design, populations, exposures or interventions, and outcome measures and chose not pool the data via meta-analysis. Thus, we have presented the findings for most outcomes narratively and in summary tables.

When statistical pooling was appropriate, this was undertaken using Review Manager (RevMan v.5.3, Copenhagen: The Nordic Cochrane Centre, the Cochrane Collaboration, 2014) via pairwise meta-analysis

When data were not presented in the format required for meta-analysis, we estimated means or standard deviations (SDs) using standard equations. We used the median instead of the mean for one study[31] for the outcomes of length of stay and operating time. Additionally, for one study[32] in the length of stay analysis where the SD could not be estimated, we substituted the mean variance of other studies within the meta-analysis.[33]

## **RESULTS**

We identified 16,083 unique records via the database searches, 56 grey literature sources, and 14 additional records in reference lists of systematic reviews. We excluded 15,016 citations by title and abstract, and another 1,090 by full text. Forty-seven studies[31, 32, 34-78] were eligible for inclusion, and 6[31, 32, 41, 58, 63, 77] were included in meta-analysis for the outcomes of operating time, intra-and post-operative complications, patient mortality and length of hospital stay. Figure 1 shows the flow of studies through the selection process.

## **Included study characteristics**

A summary of the study characteristics is provided in Table 1. Supplementary file 2 presents descriptive information for each included study. There were 45 observational studies [31, 32, 34-39, 41-74, 76-78] and two intervention studies. [40, 75] All studies were quantitative. Nearly half (n = 20/47, 43%) of the studies took place in North America, [31, 32, 35, 37, 38, 41, 45, 48, 57-60, 62, 63, 65-68, 72, 77] and

slightly more than one-third (n = 16/47, 34%) in Europe. [34, 36, 39, 40, 42, 46, 47, 50-53, 61, 64, 73, 75, JI. 76]

Table 1. Summary characteristics of the included studies

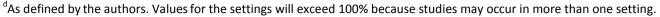
Study characteristics	n	%	Physician characteristics	n	%	Exposures, interventions and	n	%
						outcomes		
Study design			Gender			Exposures (observational) <sup>a</sup>	45	96
Cross-sectional	34	72	Reported <sup>b</sup>	38	81	Fatigue-related	15	32
Cohort	6	13	>50% male	30	79	Sleep-related	37	79
Before-after	3	6	Age			Overnight or extended shifts	18	38
RCT	2	4	Reported <sup>b</sup>	38	81	Interventions (experimental)	2	4
Time series	1	2	Range (years)	20 to	>70	Outcomes		-
Non-comparative	1	2	Specialty area <sup>c</sup>			Physician health and wellbeing	28	60
Region and country	•		Surgeons	13	28	Work and life satisfaction	9	19
North America	20	43	Anesthesiologists	10	21	Burnout	7	15
US	15	32	Generalists	7	15	Stress	8	17
Canada	4	9	ED or ICU physicians	3	6	Mental health and wellbeing	7	15
Canada, US & Mexico	1	2	Oncologists	2	4	Other health-related outcomes	5	11
Europe	16	34	Obstetrician-gynecologists	1	2	Physician performance, risk of error	21	45
France	4	9	Mixed groups	14	30	Psychomotor performance	7	15
Finland	3	6	Work setting <sup>d</sup>			Work ability and quality of care	5	11
Spain	2	4	Hospitals	37	79	Incidence of medical errors	5	11
Austria	2	4	Private practice	13	28	Surgical efficiency, effectiveness	6	13
Norway	2	4	Primary care centres, outpatient clinics	7	15	Patient outcomes	6	13
Denmark	1	2	Academic practice, training programs	5	11	1//,		-
Germany	1	2	Other (e.g., industry, military)	11	23			
Malta	1	2	Not reported	3	6			
Japan	4	9	Urban or rural					
Australia	2	4	Reported <sup>b</sup>	16	34			
Israel	2	4	Urban	12	75			
New Zealand	2	4	Rural	2	13			
United Kingdom	1	2	Mixed	2	13			

ED: emergency department; ICU: intensive care unit; RCT: randomised controlled trial; US: United States of America

<sup>&</sup>lt;sup>a</sup>Exposures that have been directly related to an outcome. Some studies included multiple exposures.

<sup>b</sup>Percentages presented using the total number of studies where the outcome was reported as the denominator.

Anesthesiologists include physician anesthetists; generalists include primary care physicians, internists, and general practitioners; mixed groups refers to studies including more than one physician group or specialty (usually large-scale surveys). In some studies, multiple distinct groups were represented.





Fifteen (32%) studies reported on fatigue-related exposures (e.g., as a source of stress, exhaustion, physical fatigue; hereafter referred to as 'fatigue'),[35, 40, 45, 48, 57, 63-71, 73] while others (n = 37/47, 79%) reported on sleep-related exposures (e.g., sleep hours, insufficient sleep, sleep deprivation, sleep disruption, sleepiness; hereafter referred to as 'insufficient sleep').[31, 32, 34, 36-47, 49-56, 58-62, 64, 67, 71, 72, 74-78] A few (n = 5/47, 11%) reported on both.[40, 45, 64, 67, 71] In some cases (n = 18/47, 38%), fatigue or insufficient sleep were related to overnight work or long on-call shifts.[31, 32, 34, 37, 40, 41, 43, 45, 46, 50, 51, 53, 58-60, 63, 75, 77] Measured outcomes varied widely and were ultimately organised into physician physical and mental health, physician performance and risk of error, and patient outcomes.

### Risk of bias appraisal

The overall quality of the body of research was poor; 62% (n = 29/47) of studies were rated at unclear or high risk of bias. Of the two randomised controlled trials, one was rated as unclear overall risk of bias[75] and one as high risk.[40] All cohort studies were at low risk of bias (mean score: 8.4/9, range: 8-9).[31, 32, 41, 58, 63, 77] All of the before-after studies were rated as high risk of bias.[34, 45, 50] The single time series study was assessed at high risk of bias.[51] The cross-sectional studies varied in performance (mean score: 3.0/5, range: 1-4); only one-third (n = 12/34, 35%) were at low risk of bias.[39, 42, 44, 47, 52, 59, 60, 69, 70, 72, 74, 76] The one non-comparative study was at unclear risk of bias.[43] Detailed assessments of the sources of bias per study are shown in Supplementary file 3.

#### Physician health and wellbeing outcomes

Twenty-eight studies reported on physician health and wellbeing-related outcomes, [35, 36, 38-40, 42, 46-48, 50-57, 60, 62, 64, 67, 68, 70-72, 74, 76, 78] including burnout (n = 7), stress (n = 8), mental health and wellbeing (n = 7), life and job satisfaction (n = 9) and other markers of health (n = 5) (Supplementary file 4).

Seven cross-sectional studies reported on burnout (5 low[39, 70, 72, 74, 76], 1 unclear[54], 1 high risk of bias[62]) among surgeons, [54, 62] anesthesiologists, [39] generalists, [76] and other mixed groups. [70, 72, 74] Two studies reported on surgeons; the larger (n = 2,564, low risk of bias) study of neurosurgeons showed increased odds of burnout with sleep deprivation (hours of sleep per night; OR 0.84, 95% CI 0.75 to 0.94, P = 0.002). [54] Among anesthesiologists one study (n = 565, low risk of bias) indicated that burnout (measured via Maslach Burnout Inventory) was more prevalent among the sleep-deprived ('lack of sleep' on one question; 47.6% vs. 16.3%, P < 0.001). [39] In one small (n = 11) study of generalists, those with burnout (measured via Pines Burnout Measure) had poorer Pittsburgh Sleep Quality Index scores (7.24±4.17 vs. 2.72±2.22, P < 0.001). [76] In the two larger studies of mixed physician groups (low risk of bias), burnout (measured via 5-point scale) was more prevalent among those who were sleep deprived (<7 hours of sleep per 24 hours; 39.6% vs. 26.4%, P < 0.05), [72] and physical fatigue ('feeling tired' on a 7-point scale) was correlated with burnout (Shirom-Melamed Burnout Measure; P = 0.88, P < 0.05). [70] In summary, evidence from 7 cross-sectional studies (71% at low risk of bias), showed associations between insufficient sleep and burnout.

Six cross-sectional studies (2 low[47, 52], 1 unclear[46], 3 high risk of bias[35, 62, 64]), one uncontrolled before-after study (high risk of bias[50]), and one intervention study (high risk of bias[40]) reported on stress outcomes among surgeons,[62] anesthesiologists,[50, 52] emergency physicians,[40, 64] internal medicine physicians, [46] and mixed groups.[35, 47] In a small sample (n = 20) of internal medicine physicians, insufficient sleep related to a 24-hour call shift showed no association with biochemical or physiological stress parameters, except levels of thyroid stimulating hormone, which was higher post-shift (P = 0.049, data not reported).[46] The remaining observational studies suggested that there was an association between insufficient sleep or fatigue and stress. The one study of orthopedic surgeons (n = 264, high risk of bias) showed that insufficient sleep (measured on a 3-point scale) and psychological distress (measured via General Health Questionnaire-12) were correlated (data not reported, P < 0.001).[62] The two reports on anesthesiologists were of varied quality; the larger (n = 328, low risk of

Seven cross-sectional studies (2 low, [52, 60] 3 unclear, [67, 71, 78] 2 high risk of bias [36, 53]) reported on aspects of mental health including addiction or substance misuse, [36, 53, 71] depression, [78] thoughts of suicide,[52] mood disturbance,[60, 71] and overall wellbeing.[67] One study,[53] which was at high risk of bias, showed no association between hours of sleep when on call and hazardous drinking behaviours (via Alcohol Use Disorder Identification Test). Meanwhile, the six other studies all showed associations between insufficient sleep and fatigue and reduced mental health. Three studies reported on anesthetists, [36, 52, 60] with two large surveys showing increased odds of tobacco (OR 1.42, 95% CI 1.04 to 1.94) and tranquilizer/hypnotics (OR 3.26, 95% CI 2.12 to 5.02) dependency being predicted by sleep deprivation (measured by one question),[36] and sleep disturbance being associated with thoughts of suicide (using a 4-point scale; P = 0.009).[52] A small study (n = 21) showed greater mood disturbance following a 17-hour night shift than a usual day (Profile of Mood States score 42.57±15.26 vs. 70.90±6.91, P < 0.001).[60] Among oncologists (n = 241), overall wellbeing was predicted by lower levels of fatigue after controlling for personal and professional characteristics (assessed via linear analog scale quality of life survey, P = 0.002).[67] A large (n = 3,862, unclear risk of bias) study of physicians showed that insufficient sleep (lower sleep hours when not at work in the past month) was associated with increased odds of depression (Quick Inventory Depressive Scale; OR 2.70, 95% CI 1.82 to 4.03 for men; OR 2.38, 95% CI 1.11 to 5.10 for women).[78] In open-ended questions, senior physicians in one study (unclear risk of bias) attributed the development of mental illness to tiredness and stress at work.[71] In summary, 7 cross-sectional studies (29% at low risk of bias) were identified, and of these 6 supported an association between insufficient sleep or fatigue and negative mental health outcomes.

Nine cross-sectional studies (4 low,[42, 47, 72, 74] 2 unclear,[55, 68] 3 high risk of bias[38, 48, 62]) reported on outcomes related to job satisfaction,[42, 47, 48, 55, 72, 74] life satisfaction,[38, 62, 72] or work-life balance.[68, 72] The six studies that investigated job satisfaction were all at low risk of bias and generally included mixed groups of physicians;[47, 72, 74] one study reported on general practitioners,[55], another on surgeons,[48] and one on mixed specialties.[42] Three studies showed that reductions in sleep duration and/or quality[47, 48, 74] were associated with reduced job satisfaction. Meanwhile one showed no association between insufficient sleep (<7 hours per 24-hour period) and career satisfaction (measured on a 5-point Likert scale),[72] and another showed no relationship between earlier sleep disturbance (Jenkins Scale) and later job demands or job control (measured via 5-point scale).[42] A single study (n = 92) reporting on rural general practitioners indicated that frequent sleep disturbance (measured on a 7-point scale) predicted the intention to retire early (OR 2.91, 95% CI 1.11 to 7.6, P < 0.05).[55] In summary, 6 cross-sectional studies (all at low risk of bias) were identified, and all but two[42, 72] of these studies showed that insufficient sleep and fatigue were associated with reductions in satisfaction.

The three studies reported on life satisfaction. [38, 62, 72] Of two studies among mixed physician groups, [38, 72] the one larger (n = 840) study showed that insufficient sleep (< 7 hours per day) was a predictor of reduced life satisfaction (measured on a 5-point Likert scale; OR 0.44, 95% CI 0.29 to 0.67, P  $\leq$  0.05). [72] One study at high risk of bias reported on orthopedic surgeons (n = 264), showing that sleep deprivation (measured via 3-point scale) was correlated with lower marital satisfaction (Revised Dyadic Adjustment Scale; data not reported, P < 0.001). [62] Two large studies at low or unclear risk of bias reported on work-life balance. [68, 72] Among oncologists (n = 1,117), reduced satisfaction with work-life balance (measured on a 5-point Likert scale) was predicted by high levels of fatigue (measured via 10-point visual analog scale), even when controlling for personal and work-related factors and burnout (OR 0.489, 95% CI 0.337 to 0.710, P < 0.001). [68] Among a mixed group of physicians (n = 840, low risk of bias), insufficient sleep (<7 hours in a typical 24-hour period) predicted a reduced perception of having balanced personal and professional commitments (5-point Likert scale; OR 0.46, 95% CI 0.31 to 0.71, P  $\leq$  0.05). [72] In summary, 3 cross-sectional studies (all unclear or high risk of bias) supported an association between insufficient sleep or fatigue and reduced life satisfaction, and 2 cross-sectional studies (50% low risk of bias) supported an association with reduced work-life balance.

Twenty-one studies reported on physician performance and safety-related outcomes, [31, 32, 34, 37, 38, 41, 43-47, 49, 50, 59, 61, 65, 66, 69, 71, 73, 75] including surgical efficiency and effectiveness (n = 6), psychomotor performance (n = 7), work ability and quality of care (n = 5) and medical errors (n = 5)(Supplementary file 4).

Four cohort studies (all low risk of bias[31, 32, 41, 63]), one before-after study (high risk of bias[34]) and one randomized controlled trial (high risk of bias[75]) examined the effects of insufficient sleep from overnight work or extended shifts, during surgeries[31, 32, 41] or laparoscopic simulations.[34, 75] We pooled the data from these studies[31, 32, 41, 63] via meta-analysis, which showed no difference in operating time (sometimes referred to as surgeon efficiency) between sleep deprived and non-sleep deprived surgeons (Figure 2; n = 50,046, MD -0.14, 95% CI -1.60 to 1.33, P = 0.86,  $I^2 = 0\%$ ). Of studies not meta-analysed, the small (n = 29) before-after study showed no impact of sleep deprivation from shiftwork nor of sleep hours on performance on a laparoscopic simulation (LapSimGyn).[34] One small (n = 64) intervention study compared a 24-hour shift to a usual work day, also finding no detriment to performance on a laparoscopic simulation (Minimally Invasive Surgical Trainer-Virtual Reality) despite

diminished sleep hours while working on-call.[75] In summary, pooled data from 4 cohort studies (100% low risk of bias) showed no effect of insufficient sleep on surgical efficiency. Additional data from one RCT (high risk of bias) and one before-after study (high risk of bias) also showed no association between insufficient sleep and performance on laparoscopic simulations.

Two before-after studies (high risk of bias[45, 50]) and five cross-sectional studies (2 low,[43, 59] 3 unclear, [37, 46] 1 high risk of bias [61]) reported on psychomotor performance outcomes among surgeons,[45] anesthesiologists,[37, 43, 50, 59] emergency physicians,[61] and internal medicine physicians.[46] Among a small group of surgeons (n = 9), performance on a virtual ring transfer task deteriorated after an on-call shift (data not reported, P < 0.05).[45] The four studies among anesthetists reported mixed findings. One small (n = 11) before-after study showed longer reaction times (690.8±73.4 vs. 746.5±113.7 milliseconds) and reduced concentration ability (26.4±23.5 vs. 56.3±23.0 on a 100-point scale, P = 0.007) following a 24-hour shift with insufficient sleep; [50] Two others found that insufficient sleep due to overnight shifts was associated with slower reaction times. [43, 59] Conversely, a small study (n = 11) found no effect of overnight shiftwork with insufficient sleep on any measure of psychomotor performance except Hopkin's Verbal Learning Test (t-score of 48.6±7.6 vs. 41.5±9.9, P = 0.04).[37] Among emergency physicians (n = 18), one study (high risk of bias) showed that those who were sleep deprived (<5 hours sleep after a 24-hour shift) had a reduced performance on most but not all psychomotor tests (Battery Test Reaction 5),[61] while among internal medicine physicians (n = 20, low risk of bias), neurocognitive parameters did not seem to worsen post-call.[46] In summary, two before-after (0% low risk of bias) and 5 cross-sectional studies (40% low risk of bias) showed mixed results for the association between fatigue or insufficient sleep and psychomotor performance.

Five cross-sectional studies (2 low,[47, 69] 1 unclear,[71] 2 high risk of bias[38, 65]) reported on associations between sleep deprivation or fatigue and work ability or perceived performance, all among mixed groups of physicians.[38, 47, 65, 69, 71] The two large studies at low risk of bias showed that sleep problems and fatigue were inversely associated with physicians' perceived quality of work.[47, 69] Among 1,541 physicians in Finland, sleeping problems (measured by 4 questions from the Jenkins Scale) were inversely associated with scores on the Work Ability Index ( $\beta$  = -0.29, P < 0.001),[47] while a study of 890 physicians from Israel demonstrated that perceived quality of care was predicted by fatigue (1 item on the Shirom-Melamed Burnout Measure) even after controlling for components of burnout ( $\beta$  = 0.17, P < 0.05).[69] Similarly, in one study, comments from senior physicians suggested that continual

tiredness and exhaustion negatively affected their perceived competence. [71] The two studies [38, 65] that were at high risk of bias had conflicting findings. In summary, 5 cross-sectional studies (40% at low risk of bias) reported on perceived work performance; those that were at low risk of bias supported an association between fatigue or insufficient sleep and reduced performance.

Five cross-sectional studies (1 low,[44] 2 unclear,[49, 66] 2 high risk of bias[38, 73]) reported on associations between insufficient sleep or fatigue and self-reported medical errors among surgeons,[66] anesthesiologists[44] and mixed groups of physicians.[38, 49, 73] A large (n = 7,905) study at unclear risk of bias showed that only 6.9% of surgeons reported fatigue as the most important contributor to medical errors.[66] Among anesthesiologists, a smaller study (n = 183) at low risk of bias showed that the risk of self-reported fatigue-related errors increased with more nights of work-related sleep disturbance (RR 1.25, 95% CI 1.06 to 1.49).[44] Two of the studies reporting on mixed groups of physicians had conflicting results,[38, 49] while another reported that physicians' opinions on the association between fatigue and prescribing errors differed by work setting.[73] One-third (34%) of community-based, 96% of hospital-based, and 8% of office-based physicians believed that there was a high or very high association between fatigue and prescribing errors (P < 0.05).[73] In summary, 5 cross-sectional studies (20% at low risk of bias) reported on self-reported errors, and these showed mixed findings for associations with fatigue or insufficient sleep.

#### **Patient Outcomes**

Six large (n = 270 to 38,978) cohort studies at low risk of bias reported on patient outcomes, all related to surgical[31, 32, 41, 58, 63, 77] or obstetric[58] procedures (Supplementary file 4). In these studies, insufficient sleep or fatigue were typically defined as overnight work prior to a daytime procedure[31, 41, 58, 63, 77]; though two studies measured sleep hours[32] or 'sleep opportunity'.[58] We pooled data for procedures performed by sleep deprived versus non-sleep deprived surgeons (or obstetrician-gynecologists in one case[58]). Analyses showed no difference in the rate of post-operative complications (Figure 3; 5 studies,[31, 32, 41, 63, 77] n = 60,201, RR 0.99, 95% CI 0.95 to 1.03, p = 0.51,  $I^2 = 0\%$ ) nor patient mortality (Figure 4; 5 studies,[31, 32, 41, 63, 77] n = 60,436, RR 0.98, 95% CI 0.84 to 1.15, p = 0.82,  $I^2 = 0\%$ ). One study[77] in the mortality analysis reported the number of deaths only as  $\le 5$ . We assumed 2 events for this study (midpoint between 0 and 5); sensitivity analysis using the lowest (i.e., 0) and highest (i.e., 5) possible number of events did not change the overall result (Supplementary file 5). We found considerable between-study heterogeneity in the analyses for intraoperative

complications ( $I^2 = 82\%$ ) and length of stay ( $I^2 = 86\%$ ), which could not be explained via subgroup analyses by procedure type, thus we have suppressed the average estimates of effect. For length of stay, the results of one study on cardiac surgeries favoured sleep deprived surgeons,[32] while the others[31,41,63] had null results. For intraoperative complications, the findings of one study[63] favoured non-sleep deprived surgeons, but the others[58,77] had null results.

#### **DISCUSSION**

Fatigue and chronic insufficient sleep are two potential drivers of reduced physician wellbeing[17, 19] that have thus far been understudied in physicians in independent practice. Burnout is becoming increasingly prevalent among physicians,[14-16] and recent research indicates that comprehensive individual- and system-level strategies are needed to address the problem.[6-9, 19, 21] We have systematically reviewed evidence from a heterogeneous array of available studies reporting on diverse outcomes related to physicians in independent practice and their patients. The included studies were often at high or unclear risk of bias, included small samples of physicians, and inconsistently measured and reported exposures and outcomes. The key message gleaned from this review is that despite growing interest in the topic of physician wellness, the robust evidence needed to inform individual and systems-level fatigue management strategies is lacking.

Traditionally, much of the fatigue-related research has focused on hazards to patients. The current review included six cohort studies showing that insufficient sleep and/or fatigue did not seem to result in increased rates of patient morality or post-operative complications; findings for length of stay and intra-operative complications were inconclusive. Evidence for psychomotor performance, surgical skills and errors suggest that there is indeed a potential for negative outcomes. The included studies, like many of the others in this and other systematic reviews,[79] employed indirect definitions that make it difficult to classify sleep deprived physicians with certainty. In recent years there has been a shift away from the singular focus on patient safety toward a more comprehensive view that also considers the detrimental effects of fatigue, sleep loss and other occupational hazards on physician wellness.[80] Evidence from this review supports that fatigue and insufficient sleep may be negatively associated with physician health and wellbeing. It is now recognized that health systems cannot be sustained by a workforce that is facing an epidemic of burnout.[19, 81, 82]

The most salient finding of this review is that the current evidence is insufficient to inform policy and practice. Correspondingly, a 2016 research summit on physician wellness and burnout outlined the need for timely, relevant and methodologically robust research to inform practice and policy.[21] The findings herein may be used as motivation for researchers and practitioners to develop and design methodologically strong research programs related to physician fatigue, inform successful research grant proposals, and lobby healthcare organizations to increase the focus on physician fatigue management programs. It will be important to make use of existing validated measures[87-89] consistently in future research. Identifying outcomes of importance to physicians and their patients should be prioritized, such that these may be collected within intervention studies. Reporting these consistently will allow for the effective synthesis of findings and reduce research waste.[90] Integrated knowledge translation strategies involving multiple stakeholder groups (e.g., physicians, patients, medical schools, physicians' associations and governing bodies, policymakers) may help to ensure that the research is relevant and facilitates decision-making.[91]

# **Strengths and Limitations**

Our systematic review is the first to synthesize evidence on the effects of fatigue and insufficient sleep on physicians in independent practice. The review is timely, given recent calls for research into individual and organisational solutions for burnout,[20, 21] and an increased focus on physician

health.[80, 81] While we have identified a diverse body of evidence, we could not draw definitive conclusions due to methodological weaknesses (e.g., 62% at high risk of bias, reliance primarily on cross-sectional designs and uncontrolled studies, subjective measurement of exposures and outcomes, small sample sizes, inclusion of predominantly male physicians within urban settings) and heterogeneous outcome measures in the included studies. Given that the 2017 update search was limited to one database, it is possible that a small number of relevant studies could have been missed. We believe that the likelihood that these might alter the conclusions of the review is low. The findings may have been influenced by publication bias, and may not be generalized to all settings, given our restriction to high income countries.

#### **CONCLUSION**

The evidence synthesized in this review suggests that fatigue and insufficient sleep are associated with some detrimental physician health and wellbeing outcomes; the evidence for potential associations with performance and safety outcomes was mixed. Meta-analyses for patient outcomes demonstrated that in many cases, potential relationships with physician sleep deprivation remain unclear. Our overall confidence in the findings is low, owing to a body of research that is hindered by methodological weaknesses. Further methodologically robust research that includes consistent outcomes that are of interest to physicians and their patients is needed to inform strong practice recommendations and policy decisions.

#### **ACKNOWLEDGMENTS**

We would like to thank Dr. Meghan Sebastianski and Andrea Milne for assistance with screening, and Dr. Allison Gates for screening, data verification, and quality assessment of the French studies.

# **COPYRIGHT/LICENSE FOR PUBLICATION**

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide license to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the

#### **COMPETING INTERESTS**

All authors have completed the ICMJE uniform disclosure form at <a href="www.icmje.org/coi\_disclosure.pdf">www.icmje.org/coi\_disclosure.pdf</a> and declare: Dr. Christopher Simon is employed by the Canadian Medical Association, who provided financial support for the research; there are no other relationships or activities that could appear to have influenced the submitted work.

#### **CONTRIBUTOR STATEMENT**

All authors contributed to the conception and design of the project. MG and AW contributed to the acquisition, analysis and interpretation of the data, and drafted the manuscript. RF contributed to acquisition of data. CSa, CSi and MPD contributed to interpretation of data and revised the manuscript for important intellectual content. All authors approved the final version of the manuscript as submitted.

### **FUNDING STATEMENT**

This research was supported by the Canadian Medical Association. The funder had no role in the study design; collection, analysis and interpretation of data; the writing of the report; and the decision to submit the article for publication.

#### **ROLE OF FUNDERS**

Dr. Christopher Simon is employed by the Canadian Medical Association. The remaining authors are independent from the funders. The funders had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

# TRANSPARENCY DECLARATION

The lead author (MG) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; no important aspects of the study have been omitted; and all discrepancies from the study as planned have been explained.

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

# **DATA ACCESS STATEMENT**

All authors, external and internal, had full access to all of the data in the study and can take responsibility for the integrity of the data and the accuracy of the interpretation.

#### **DATA SHARING STATEMENT**

The data pertaining to this systematic review are available from the corresponding author upon reasonable request.

#### **FIGURE CAPTIONS**

- Figure 1. Flow of records through the selection process
- Figure 2. Forest plot for operating time among sleep deprived and non-sleep deprived surgeons
- **Figure 3.** Forest plot for post-operative complications among surgeries performed by sleep deprived and non-sleep deprived surgeons
- **Legend:** Vinden 2013 reported iatrogenic injuries; Schieman 2008, Govindarajan 2015, and Chu 2011 reported post-operative complication rate; Ellman 2004 reported post-operative complications (other types of complications reported not included in the analysis)
- **Figure 4.** Forest plot for patient mortality among surgeries performed by sleep deprived and non-sleep deprived surgeons

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

# 1. Greig P, Snow R. Fatigue and risk: are train drivers safer than doctors? *BMJ* 2017;359. doi: https://doi.org/10.1136/bmj.j5107

- 2. Parker JBR. The effects of fatigue on physician performance an underestimated cause of physician impairment and increased patient risk. *Can J Anaesth* 1987;34(5):489-95.
- 3. Temple J. Resident duty hours around the globe: where are we now? BMC Med Educ 2014;14(Suppl 1):S8. https://doi.org/10.1186/1472-6920-14-S1-S8
- 4. Imrie K, Frank J, The National Steering Committee on Resient Duty Hours. Fatigue, risk, & excellence: towards a pan-Canadian consensus on resident duty hours. Ottawa, Ontario: The Royal College of Physicians and Surgeons of Canada, 2013. <a href="http://www.residentdutyhours.ca/final\_report.php">http://www.residentdutyhours.ca/final\_report.php</a> (accessed 12 Jan 2018).
- Accreditation Council for Graduate Medical Education. History of duty hours. 2017.
   <a href="http://www.acgme.org/What-We-Do/Accreditation/Clinical-Experience-and-Education-formerly-Duty-Hours/History-of-Duty-Hours">http://www.acgme.org/What-We-Do/Accreditation/Clinical-Experience-and-Education-formerly-Duty-Hours/History-of-Duty-Hours</a> (accessed 12 Jan 2018).
- 6. Bolster L, Rourke L. The effect of restricting residents' duty hours on patient safety, resident well-being, and resident education: an updated systematic review. *J Grad Med Educ* 2015;7(3):349-63.
- 7. Harris JD, Staheli G, LeClere L, Andersone D, McCormick F. What effects have resident work-hour changes had on education, quality of life, and safety? A systematic review. *Clin Orthop* 2015;473(5):1600-8.
- 8. Ahmed N, Devitt KS, Keshet I, et al. A systematic review of the effects of resident duty hour restrictions in surgery: impact on resident wellness, training, and patient outcomes. *Ann Surg* 2014;259(6):1041-53.
- 9. Peets A, Ayas NT. Restricting resident work hours: the good, the bad, and the ugly. *Crit Care Med* 2012;40(3):960-6.
- Canadian Medical Association. CMA Policy: Management of physician fatigue. Ottawa, Canada:
   Canadian Medical Association; 2014.
- 11. Anim M, Markert RJ, Wood VC, Schuster BL. Physician practice patterns resemble ACGME duty hours. *Am J Med* 2009;122:587-93.
- 12. Transport Canada. Fatigue risk management system for the canadian aviation industry: policies and procedures development guidelines. Ottawa, Canada: Her Majesty the Queen in Right of Canada, as

https://www.tc.gc.ca/eng/civilaviation/publications/TP14576-6042.htm (accessed 12 Jan 2018).

- 13. Patterson PD, Higgins JS, Lang ES, et al. Evidence-based guidelines for fatigue risk management in ems: formulating research questions and selecting outcomes. *Prehosp Emerg Care* 2017;21(2):149-56.
- 14. Boudreau RA, Grieco RL, Cahoon SL, Robertson RC, Wedel RJ. The pandemic from within: two surveys of physician burnout in Canada. *Can J Commun Ment Health* 2006;25(2):71-88.
- 15. Kumar S. Burnout and doctors: prevalence, prevention and intervention. *Healthcare* 2016;4(3):37. doi: 10.3390/healthcare4030037
- 16. Shanafelt TD, Boone S, Tan L, et al. Burnout and satisfaction with work-life balance among us physicians relative to the general us population. *Arch Intern Med* 2012;172(18):1377-85.
- 17. Brady KJS, Trockel MT, Khan CT, et al. What do we mean by physician wellness? A systematic review of its definition and measurement. *Acad Psychiatry* 2017. doi: 10.1007/s40596-017-0781-6
- 18. Fralick M, Flegel K. Physician burnout: who will protect us from ourselves? *CMAJ* 2014;186(10):731. doi: 10.1503/cmaj.140588
- 19. Shanafelt TD, Noseworthy JH. Executive leadership and physician well-being: nine organizational strategies to promote engagement and reduce burnout. *Mayo Clin Proc* 2017;92(1):129-46.
- 20. Lemaire JB, Wallace JE. Burnout among doctors. BMJ 2017;358:j3360. doi: 10.1136/bmj.j3360
- 21. Dyrbye LN, Trockel M, Frank E, et al. Development of a research agenda to identify evidence-based strategies to improve physician wellness and reduce burnout. *Ann Intern Med* 2017;166(10):743-4.
- 22. Higgins JPT, Green S (editors). The Cochrane handbook for systematic reviews of interventions, version 5.1.0. London, UK: The Cochrane Collaboration, 2011. <a href="http://www.handbook.cochrane.org">http://www.handbook.cochrane.org</a> (accessed 12 Jan 2018).
- 23. 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(7):e1000100. https://doi.org/10.1371/journal.pmed.1000100
- 24. Pattani R, Wu PE, Dhalla IA. Resident duty hours in Canada: past, present and future. *CMAJ* 2014;186(10):761-5.
- 25. International Monetary Fund. World economic outlook: too slow for too long. Washington, DC: International Monetary Fund, 2016. <a href="http://www.imf.org/external/pubs/ft/weo/2016/01/">http://www.imf.org/external/pubs/ft/weo/2016/01/</a> (accessed 12 Jan 2018).

- 27. Wells GA SB, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa, Canada: The Ottawa Hospital Research Institute, 2014.
  <a href="http://www.ohri.ca/programs/clinical\_epidemiology/oxford.asp">http://www.ohri.ca/programs/clinical\_epidemiology/oxford.asp</a> (accessed 12 Jan 2018).
- 28. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-88.
- 29. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-60.
- 30. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315(7109):629-34.
- 31. Govindarajan A, Urbach DR, Kumar M, et al. Outcomes of daytime procedures performed by attending surgeons after night work. *N Engl J Med* 2015;373(9):845-53.
- 32. Chu MW, Stitt LW, Fox SA, Kiaii B, Quantz M, Guo L, et al. Prospective evaluation of consultant surgeon sleep deprivation and outcomes in more than 4000 consecutive cardiac surgical procedures. *Arch Surg* 2011;146(9):1080-5.
- 33. Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *J Clin Epidemiol* 59(1):7-10.
- 34. Amirian I, Andersen LT, Rosenberg J, Gogenur I. Laparoscopic skills and cognitive function are not affected in surgeons during a night shift. *J Surg Educ* 2014;71(4):543-50.
- 35. Aziz A. Sources of perceived stress among american medical doctors: a cross-cultural perspective. Cross Cultural Management 2004;11(4):28-39.
- 36. Beaujouan L, Czernichow S, Pourriat JL, Bonnet F. Prevalence and risk factors for substance abuse and dependence among anaesthetists: a national survey. *Ann Fr Anesth Réanim* 2005;24(5):471-9.
- 37. Chang LC, Mahoney JJ, 3rd, Raty SR, Ortiz J, Apodaca S, De La Garza R 2nd. Neurocognitive effects following an overnight call shift on faculty anesthesiologists. *Acta Anaesthesiol Scand* 2013;57(8):1051-7.
- 38. Chen I, Vorona R, Chiu R, Ware JC. A survey of subjective sleepiness and consequences in attending physicians. *Behav Sleep Med* 2008;6(1):1-15.

- 39. Doppia MA, Estryn-Béhar M, Fry C, Guetarni K, Lieutaud T, comité de pilotage de l'enquête SESMAT. Burnout in French doctors: a comparative study among anaesthesiologists and other specialists in French hospitals (SESMAT study). *Ann Fr Anesth Réanim* 2011;30(11):782-94.
- 40. Dutheil F, Trousselard M, Perrier C, et al. Urinary interleukin-8 is a biomarker of stress in emergency physicians, especially with advancing age--the JOBSTRESS randomized trial. *PLoS One* 2013;8(8):e71658. https://doi.org/10.1371/journal.pone.0071658
- 41. Ellman PI, Law MG, Tache-Leon C, et al. Sleep deprivation does not affect operative results in cardiac surgery. *Ann Thorac Surg* 2004;78(3):906-11.
- 42. Elovainio M, Heponiemi T, Jokela M, et al. Stressful work environment and wellbeing: what comes first? *J Occup Health Psychol* 2015;20(3):289-300.
- 43. Gander P, Millar M, Webster C, Merry A. Sleep loss and performance of anaesthesia trainees and specialists. *Chronobiol Int* 2008;25(6):1077-91.
- 44. Gander PH, Merry A, Millar MM, Weller J. Hours of work and fatigue-related error: a survey of New Zealand anaesthetists. *Anaesth Intensive Care* 2000;28(2):178-83.
- 45. 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(6):813-9.
- 46. Harbeck B, Suefke S, Haas CS, Lehnert H, Kropp P, Moenig H. No stress after 24-hour on-call shifts? *J Occup Health* 2015;57(5):438-47.
- 47. Heponiemi T, Puttonen S, Elovainio M. On-call work and physicians' well-being: testing the potential mediators. *Occup Med* 2014;64(5):352-7.
- 48. Jackson TN, Pearcy CP, Khorgami Z, Agrawal V, Taubman KE, Truitt MS. The physician attrition crisis: a cross-sectional survey of the risk factors for reduced job satisfaction among US surgeons. *World J Surg* 2017;24. doi: 10.1007/s00268-017-4286-y
- 49. Kaneita Y, Ohida T. Association of current work and sleep situations with excessive daytime sleepiness and medical incidents among Japanese physicians. *J Clin Sleep Med* 2011;7(5):512-22.
- 50. Lederer W, Kopp M, Hahn O, et al. Post-duty psychomotor performance in young and senior anaesthetists. *Eur J Anaesthesiol* 2006;23(3):251-6.
- 51. Leichtfried V, Putzer G, Perkhofer D, Schobersberger W, Benzer A. Circadian melatonin profiles during single 24-h shifts in anesthetists. *Sleep Breath* 2011;15(3):503-12.
- 52. Lindfors PM, Nurmi KE, Meretoja OA, et al. On-call stress among Finnish anaesthetists. *Anaesthesia* 2006;61(9):856-66.

- 54. Nishimura K, Nakamura F, Takegami M, et al. Cross-sectional survey of workload and burnout among Japanese physicians working in stroke care: the nationwide survey of acute stroke care capacity for proper designation of comprehensive stroke center in Japan (J-ASPECT) study. *Circ Cardiovasc Qual Outcomes* 2014;7(3):414-22.
- 55. Pit SW, Hansen V. Factors influencing early retirement intentions in Australian rural general practitioners. *Occup Med* 2014;64(4):297-304.
- 56. Pit SW, Hansen V. The relationship between lifestyle, occupational health, and work-related factors with presenteeism amongst general practitioners. *Arch Environ Occup Health* 2016;71(1):49-56.
- 57. Roberts DL, Shanafelt TD, Dyrbye LN, West CP. A national comparison of burnout and work-life balance among internal medicine hospitalists and outpatient general internists. *J Hosp Med* 2014;9(3):176-81.
- 58. Rothschild JM, Keohane CA, Rogers S, et al. Risks of complications by attending physicians after performing nighttime procedures. *JAMA* 2009;302(14):1565-72.
- 59. Saadat H, Bissonnette B, Tumin D, et al. Effects of partial sleep deprivation on reaction time in anesthesiologists. *Paediatr Anaesth* 2017;27(4):358-62.
- 60. Saadat H, Bissonnette B, Tumin D, et al. Time to talk about work-hour impact on anesthesiologists: The effects of sleep deprivation on Profile of Mood States and cognitive tasks. *Paediatr Anaesth* 2016;26(1):66-71.
- 61. Sanches I, Teixeira F, dos Santos JM, Ferreira AJ. Effects of acute sleep deprivation resulting from night shift work on young doctors. *Acta Med Port* 2015;28(4):457-62.
- 62. Sargent MC, Sotile W, Sotile MO, Rubash H, Barrack RL. Quality of life during orthopaedic training and academic practice. Part 1: orthopaedic surgery residents and faculty. *J Bone Joint Surg Am* 2009;91(10):2395-405.
- 63. Schieman C, MacLean AR, Buie WD, Rudmik LR, Ghali WA, Dixon E. Does surgeon fatigue influence outcomes after anterior resection for rectal cancer? *Am J Surg* 2008;195(5):684-7.
- 64. Sende J, Jbeili C, Schvahn S, Khalid M, Asaph J, Romano H, et al. Stress factors and stress consequences among emergency physicians: vational survey. *Annales Françaises de Médecine d'Urgence* 2012;2(4):224-31.

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

- 65. Sexton JB, Thomas EJ, Helmreich RL. Error, stress, and teamwork in medicine and aviation: Cross sectional surveys. *Hum Perform Extrem Environ* 2001;6(1):6-11.
- 66. Shanafelt TD, Balch CM, Bechamps G, et al. Burnout and medical errors among American surgeons. *Ann Surg* 2010;251(6):995-1000.
- 67. Shanafelt TD, Novotny P, Johnson ME, et al. The well-being and personal wellness promotion strategies of medical oncologists in the North Central Cancer Treatment Group. *Oncology* 2005;68(1):23-32.
- 68. Shanafelt TD, Raymond M, Kosty M, et al. Satisfaction with work-life balance and the career and retirement plans of US oncologists. *J Clin Oncol* 2014;32(11):1127-35.
- 69. Shirom A, Nirel N, Vinokur AD. Overload, autonomy, and burnout as predictors of physicians' quality of care. *J Occup Health Psychol* 2006;11(4):328-42.
- 70. Shirom A, Nirel N, Vinokur AD. Work hours and caseload as predictors of physician burnout: the mediating effects by perceived workload and by autonomy. *J Appl Psychol* 2010;59(4):539-65.
- 71. Smith F, Goldacre MJ, Lambert TW. Adverse effects on health and wellbeing of working as a doctor: views of the UK medical graduates of 1974 and 1977 surveyed in 2014. *J R Soc Med* 2017;110(5):198-207.
- 72. Starmer AJ, Frintner MP, Freed GL. Work-life balance, burnout, and satisfaction of early career pediatricians. *Pediatrics* 2016;e20153183. doi: 10.1542/peds.2015-3183
- 73. Tanti A, Camilleri M, Borg AA, et al. Opinions of Maltese doctors and pharmacists on medication errors. *Int J Saf Med* 2017;29(1-2):81-99.
- 74. Tokuda Y, Hayano K, Ozaki M, Bito S, Yanai H, Koizumi S. The interrelationships between working conditions, job satisfaction, burnout and mental health among hospital physicians in Japan: a path analysis. *Ind Health* 2009;47(2):166-72.
- 75. 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(6):753-7.
- 76. Vela-Bueno A, Moreno-Jimenez B, Rodriguez-Munoz A, et al. Insomnia and sleep quality among primary care physicians with low and high burnout levels. *J Psychosom Res* 2008;64(4):435-42.
- 77. Vinden C, Nash DM, Rangrej J, et al. Complications of daytime elective laparoscopic cholecystectomies performed by surgeons who operated the night before. *Obstet Gynecol Surv* 2014;69(2):71-3.

- 79. Sturm L, Dawson D, Vaughan R, et al. Effects of fatigue on surgeon performance and surgical outcomes: a systematic review. *ANZ J Surg* 2011;81(7-8):502-9.
- 80. Wallace JE, Lemaire JB, Ghali WA. Physician wellness: a missing quality indicator. *Lancet* 2009;374(9702):1714-21.

- 81. Canadian Medical Association. CMA policy: physician health. Ottawa, Canada: Canadian Medical Assocation, 2017. <a href="http://policybase.cma.ca/dbtw-wpd/Policypdf/PD18-01.pdf">http://policybase.cma.ca/dbtw-wpd/Policypdf/PD18-01.pdf</a> (accessed 12 Jan 2018).
- 82. Shanafelt T, Goh J, Sinsky C. The business case for investing in physician well-being. *JAMA Intern Med* 2017;177(2):195-205.
- 83. Gaba DM, Howard SK. Fatigue among clinicians and the safety of patients. *N Engl J Med* 2002;347(16):1249-55.
- 84. Ward S, Outram S. Medicine: in need of culture change. Itern Med J 2016;46(1):112-6.
- 85. Panagioti M, Panagopoulou E, Bower P, et al. Controlled interventions to reduce burnout in physicians: a systematic review and meta-analysis. *JAMA Intern Med* 2017;177(2):195-205.
- 86. American Medical Association. STEPSforward<sup>TM</sup>. Chicago, IL: American Medical Association, 2017. https://www.stepsforward.org/modules/joy-in-medicine (accessed 12 Jan 2018).
- 87. Jackson C. The Chalder Fatigue scale (CFQ 11). Occup Med 2015;65(1):86. doi: 10.1093/occmed/kgu168
- 88. Chalder T, Berelowitz G, Pawlikowska T, et al. Development of a fatigue scale. *J Psychosom Res* 1993;37(2):147-53.
- 89. Smets E, Garssen B, Bonke Bd, De Haes J. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *J Psychosom Res* 1995;39(3):315-25.
- 90. Ioannidis JPA, Greenland S, Hlatky MA, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;383(9912):166-75.
- 91. Gagliardi AR, Berta W, Kothari A, Boyko J, Urquhart R. Integrated knowledge translation (IKT) in health care: a scoping review. *Implement Sci* 2016;11:38. https://doi.org/10.1186/s13012-016-0399-1

To be exterior only

BMJ Open: first published as 10.1136/bmjopen-2018-021967 on 21 September 2018. Downloaded from http://bmjopen.bmj.com/ on June 9, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

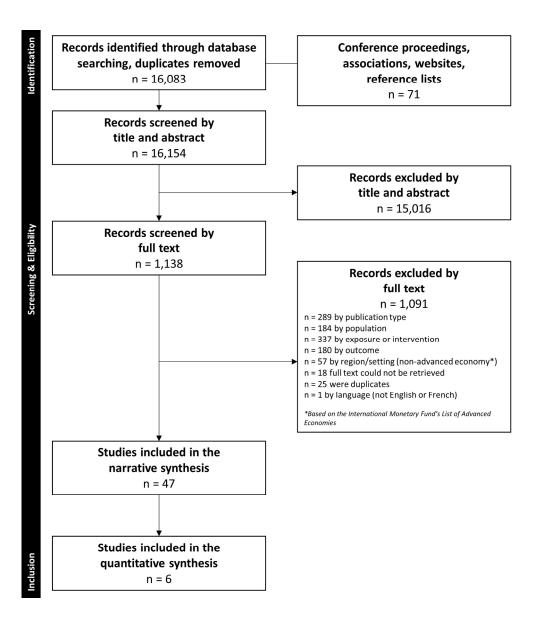


Figure 1. Flow of records through the selection process  $190x215mm (300 \times 300 DPI)$ 

	Sle	ep depriv	ed	Not sl	eep depi	rived		Mean Difference		Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	Year	IV, Random, 95% CI
Ellman 2004	107.7	55.24	339	107.4	56.05	6412	5.9%	0.30 [-5.74, 6.34]	2004	+
Schieman 2008	114	226.6	22	148	226.6	248	0.0%	-34.00 [-132.80, 64.80]	2008	<del></del>
Chu 2011	110.2	119.13	1678	114	120.88	2369	3.8%	-3.80 [-11.30, 3.70]	2011	<u>+</u>
Govindarajan 2015	156	75.56	19489	156	80	19489	90.2%	0.00 [-1.54, 1.54]	2015	
Total (95% CI) Heterogeneity: Tau* = Test for overall effect:				P = 0.70	); I*= 0%		100.0%	-0.14 [-1.60, 1.33]		-100 -50 0 50 100 Favors sleep deprived Favors non-sleep deprived

Figure 2. Forest plot for operating time among sleep deprived and non-sleep deprived surgeons



	Sleep de	prived	Not sleep d	eprived		Risk Ratio	Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI	M-H, Random, 95% CI
Chu 2011	259	1678	404	2369	7.8%	0.91 [0.78, 1.04]	· ·
Ellman 2004	29	339	480	6412	1.2%	1.14 [0.80, 1.64]	<u> </u>
Govindarajan 2015	3527	19489	3543	19489	89.4%	1.00 [0.95, 1.04]	
Schieman 2008	12	22	164	248	1.0%	0.82 [0.56, 1.22]	<del></del>
Vinden 2014	14	2031	72	8124	0.5%	0.78 [0.44, 1.38]	
Total (95% CI)		23559		36642	100.0%	0.99 [0.95, 1.03]	ı •
Total events	3841		4663				
Heterogeneity: Tau <sup>2</sup> =	0.00; Chi2	= 3.69, d	f = 4 (P = 0.45	5); I <sup>2</sup> = 0%			
Test for overall effect:	Z = 0.66 (F	P = 0.51)					0.2 0.5 1 2 5 Favors sleep deprived Favors non-sleep deprived

Figure 3. Forest plot for post-operative complications among surgeries performed by sleep deprived and non-sleep deprived surgeons

Legend: Vinden 2013 reported iatrogenic injuries; Schieman 2008, Govindarajan 2015, and Chu 2011 reported post-operative complication rate; Ellman 2004 reported post-operative complications (other types of complications reported not included in the analysis)



Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

Figure 4. Forest plot for patient mortality among surgeries performed by sleep deprived and non-sleep deprived surgeons

381x101mm (300 x 300 DPI)

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text and

Records retrieved: 5068 and 1442 in the update (removed duplicates retrieved in previous search)

- 1. Medical Staff, Hospital/
- 2. Physician Impairment/
- 3. exp Physicians/
- 4. allergist\*.ti.

- 5. (an?esthetist\* or an?esthesiologist\*).ti.
- 6. cardiologist\*.ti.
- 7. clinician\*.ti.
- 8. (clinician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 9. (clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw,kf.
- 10. dermatologist\*.ti.
- 11. endocrinologist\*.ti.
- 12. doctor\*.ti.
- 13. (doctor\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 14. (doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw,kf.
- 15. family practitioner\*.ti.
- 16. gastroenterologist\*.ti.
- 17. (general practitioner\* or GP\*).ti.
- 18. (general adj2 physician\*).ti.
- 19. geriatrician\*.ti.
- 20. gyn?ecologist\*.ti.

```
21. h?ematologist*.ti.
```

- 22. (health\* adj2 (professional\* or provider\*)).ti.
- 23. hospitalist\*.ti.
- 24. (house staff\* or housestaff\*).ti.
- 25. intensivist\*.ti.
- 26. internist\*.ti.
- 27. medical professional\*.ti.
- 28. obstetrician\*.ti.
- 29. oncologist\*.ti.
- 30. ophthalmologist\*.ti.
- 31. orthop?edist\*.ti.
- 32. (otolaryngologist\* or otorhinolaryngologist\*).ti.
- 33. neonatologist\*.ti.
- 34. nephrologist\*.ti.
- 35. neurologist\*.ti.
- 36. neuropsychiatrist\*.ti.
- 37. neurosurgeon\*.ti.
- 38. p?ediatrician\*.ti.
- 39. perinatologist\*.ti.
- 40. physician\*.ti.
- 41. (physician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 42. (physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw,kf.
- 43. primary care practitioner\*.ti.
- 44. psychiatrist\*.ti.
- 45. pulmonologist\*.ti.
- 46. rheumatologist\*.ti.
- 47. surgeon\*.ti.

- 48. (surgeon\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 49. traumatologist\*.ti.
- 50. urologist\*.ti.
- 51. or/1-50 [Combined MeSH, title, and text word searches for physicians]
- 52. Burnout, Professional/
- 53. exp Circadian Rhythm/
- 54. exp Fatigue/
- 55. Occupational Health/
- 56. Rest/ph, px [Physiology, Psychology]
- 57. Sleep Deprivation/
- 58. Sleep Disorders, Circadian Rhythm/
- 59. Sleep Wake Disorders/
- 60. exp Stress, Psychological/
- 61. Workload/px [Psychology]
- 62. Work Schedule Tolerance/
- 63. ((24 hour\* or 24 hr\* or twenty four hour\* or twentyfour hour\*) adj rhythm\*).tw,kf.
- 64. biological rhythm\*.tw,kf.
- 65. (burn out\* or burned out\* or burnt out\* or burnout\*).tw,kf.
- 66. circadian misalignment.tw,kf.
- 67. ((circadian or diurnam or ultradian) adj rhythm\*).tw,kf.
- 68. exhaust\*.tw,kf.
- 69. fatigu\*.tw,kf.
- 70. (sleep\* adj3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)).tw,kf.
- 71. tired\*.tw,kf.
- 72. weariness.tw,kf.
- 73. or/52-72 [Combined MeSH and text words for fatigue]
- 74. and/51,73 [Combined concepts for physicians and fatigue]
- 75. animals/ not (animals/ and humans/)
- 76. 74 not 75
- 77. (comment or editorial or letter).pt.

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

```
78. 76 not 77
```

- 79. limit 78 to yr="2000-Current"
- 80. limit 79 to (english or french)
- 81. remove duplicates from 80

Database: Ovid Embase 1996 to 2016 Week 15

Date searched: 13 April 2016

Records retrieved: 8859

- 1. medical staff/
- 2. exp physician/
- 3. allergist\*.ti.
- 4. (an?esthetist\* or an?esthesiologist\*).ti.
- 5. cardiologist\*.ti.
- 6. clinician\*.ti.
- 7. (clinician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 8. (clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 9. dermatologist\*.ti.
- 10. endocrinologist\*.ti.
- 11. doctor\*.ti.
- 12. (doctor\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 13. (doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 14. family practitioner\*.ti.
- 15. gastroenterologist\*.ti.
- 16. (general practitioner\* or GP\*).ti.
- 17. (general adj2 physician\*).ti.

- 19. gyn?ecologist\*.ti.
- 20. h?ematologist\*.ti.
- 21. (health\* adj2 (professional\* or provider\*)).ti.
- 22. hospitalist\*.ti.
- 23. (house staff\* or housestaff\*).ti.
- 24. intensivist\*.ti.
- 25. internist\*.ti.
- 26. medical professional\*.ti.
- 27. obstetrician\*.ti.
- 28. oncologist\*.ti.
- 29. ophthalmologist\*.ti.
- 30. orthop?edist\*.ti.
- 31. (otolaryngologist\* or otorhinolaryngologist\*).ti.
- 32. neonatologist\*.ti.
- 33. nephrologist\*.ti.
- 34. neurologist\*.ti.
- 35. neuropsychiatrist\*.ti.
- 36. neurosurgeon\*.ti.
- 37. p?ediatrician\*.ti.
- 38. perinatologist\*.ti.
- 39. physician\*.ti.
- 40. (physician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 41. (physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 42. primary care practitioner\*.ti.
- 43. psychiatrist\*.ti.
- 44. pulmonologist\*.ti.
- 45. rheumatologist\*.ti.
- 46. surgeon\*.ti.

- 47. (surgeon\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 48. traumatologist\*.ti.
- 49. urologist\*.ti.
- 50. or/1-49 [Combined Emtree, title, and text word searches for physicians]
- 51. burnout/
- 52. circadian rhythm/
- 53. circadian rhythm sleep disorder/
- 54. fatigue/
- 55. mental stress/
- 56. occupational health/
- 57. sleep deprivation/
- 58. sleep waking cycle/
- 59. work capacity/
- 60. work schedule/
- 61. working time/
- 62. workload/
- 63. ((24 hour\* or 24 hr\* or twenty four hour\* or twentyfour hour\*) adj rhythm\*).tw.
- 64. biological rhythm\*.tw.
- 65. (burn out\* or burned out\* or burnt out\* or burnout\*).tw.
- 66. circadian misalignment.tw.
- 67. ((circadian or diurnam or ultradian) adj rhythm\*).tw.
- 68. exhaust\*.tw.
- 69. fatigu\*.tw.
- 70. (sleep\* adj3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)).tw.
- 71. tired\*.tw.
- 72. weariness.tw.
- 73. or/51-72 [Combined Emtree and text words for fatigue]
- 74. and/50,73 [Combined concepts for physicians and fatigue]
- 75. animals/ not (animals/ and humans/)
- 76. 74 not 75

78. 76 not 77

- 79. limit 78 to yr="2000-Current"
- 80. limit 79 to (english or french)
- 81. limit 80 to embase

Database: Ovid PsycINFO 1987 to April Week 1 2016

Date searched: 13 April 2016

Records retrieved: 2094

- 1. exp Physicians/
- 2. allergist\*.ti.
- 3. (an?esthetist\* or an?esthesiologist\*).ti.
- 4. cardiologist\*.ti.
- 5. clinician\*.ti.
- 6. (clinician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 7. (clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 8. dermatologist\*.ti.
- 9. endocrinologist\*.ti.
- 10. doctor\*.ti.
- 11. (doctor\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 12. (doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 13. family practitioner\*.ti.
- 14. gastroenterologist\*.ti.
- 15. (general practitioner\* or GP\*).ti.
- 16. (general adj2 physician\*).ti.

- 17. geriatrician\*.ti.
- 18. gyn?ecologist\*.ti.
- 19. h?ematologist\*.ti.
- 20. (health\* adj2 (professional\* or provider\*)).ti.
- 21. hospitalist\*.ti.
- 22. intensivist\*.ti.
- 23. internist\*.ti.
- 24. medical professional\*.ti.
- 25. obstetrician\*.ti.
- 26. oncologist\*.ti.
- 27. ophthalmologist\*.ti.
- 28. orthop?edist\*.ti.
- 29. (otolaryngologist\* or otorhinolaryngologist\*).ti.
- 30. neonatologist\*.ti.
- 31. nephrologist\*.ti.
- 32. neurologist\*.ti.
- 33. neuropsychiatrist\*.ti.
- 34. neurosurgeon\*.ti.
- 35. p?ediatrician\*.ti.
- 36. perinatologist\*.ti.
- 37. physician\*.ti.
- 38. (physician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 39. (physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 40. primary care practitioner\*.ti.
- 41. psychiatrist\*.ti.
- 42. pulmonologist\*.ti.
- 43. rheumatologist\*.ti.
- 44. surgeon\*.ti.

- 45. (surgeon\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 46. traumatologist\*.ti.
- 47. urologist\*.ti.
- 48. or/1-47 [Combined thesaurus, title, and text word searches for physicians]
- 49. Compassion Fatigue/
- 50. Fatigue/
- 51. Human Biological Rhythms/
- 52. Occupational Health/
- 53. Occupational Stress/
- 54. Sleep/
- 55. Sleepiness/
- 56. Working Conditions/
- 57. Work Rest Cycles/
- 58. Work Week Length/
- 59. Work Scheduling/
- 60. Workday Shifts/
- 61. ((24 hour\* or 24 hr\* or twenty four hour\* or twentyfour hour\*) adj rhythm\*).tw.
- 62. biological rhythm\*.tw.
- 63. (burn out\* or burned out\* or burnt out\* or burnout\*).tw.
- 64. circadian misalignment.tw.
- 65. ((circadian or diurnam or ultradian) adj rhythm\*).tw.
- 66. exhaust\*.tw.
- 67. fatigu\*.tw.
- 68. (sleep\* adj3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)).tw.
- 69. tired\*.tw.
- 70. weariness.tw.
- 71. or/49-70 [Combined thesaurus and text words for fatigue]
- 72. and/48,71 [Combined concepts for physicians and fatigue]
- 73. limit 72 to yr="2000-Current"
- 74. limit 73 to (english or french)

Database: CINAHL Plus with Full Text (1937 to the present) via EBSCOhost

Date searched: 14 April 2016

Records retrieved: 3378

- S1. (MH "Medical Staff, Hospital+")
- S2. (MH "Physicians+")
- S3. TI allertist\*
- S4. TI (anesthetist\* or anaesthetist\* or anesthesiologist\* or anaesthesiologist\*)
- S5. TI cardiologist\*
- S6. TI clinician\*
- S7. clinician\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or "work\* hour\*" or "work life balance")
- S8. clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)
- S9. TI dermatologist\*
- S10. TI endocrinologist\*
- S11. TI doctor\*
- S12. doctor\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or work\* hour\* or "work life balance")
- S13. doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)
- S14. TI "family practitioner\*"
- S15. TI gastroenterologist\*
- S16. TI ("general practitioner\*" or GP\*)
- S17. TI (general N2 physician\*)
- S18. TI geriatrician\*
- S19. TI (gynaecologist\* or gynecologist\*)
- S20. TI (haematologist\* or hematologist\*)
- S21. TI hospitalist\*
- S22. TI ("house staff\*" or housestaff\*)

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

- S23. TI intensivist\*
- S24. TI internist\*

- S25. TI obstetrician\*
- S26. TI oncologist\*
- S27. TI ophthalmologist\*
- S28. TI (orthopaedist\* or orthopedist\*)
- S29. TI (otolaryngologist\* or otorhinolaryngologist\*)
- S30. TI neonatologist\*
- S31. TI nephrologist\*
- S32. TI neurologist\*
- S33. TI neuropsychiatrist\*
- S34. TI neurosurgeon\*
- S35. TI (paediatrician\* OR pediatrician\*)
- S36. TI perinatologist\*
- S37. TI physician\*
- S38. physician\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or "work\* hour\*" or "work life balance")
- S39. physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)
- S40. TI "primary care practitioner\*"
- S41. TI psychiatrist\*
- S42. TI pulmonologist\*
- S43. TI rheumatologist\*
- S44. TI surgeon\*
- S45. surgeon\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or work\* hour\* or "work life balance")
- S46. TI traumatologist\*
- S47. TI urologist\*
- S48. S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

S42 OR S43 OR S44 OR S45 OR S46 OR S47

- S49. (MH "Circadian Rhythm")
- S50. (MH "Fatigue")
- S51. (MH "Impairment, Health Professional")
- S52. (MH "Mental Fatigue")
- S53. (MH "Occupational Health")
- S54. (MH "Shiftwork")
- S55. (MH "Sleep Deprivation")
- S56. (MH "Sleep Disorders, Circadian Rhythm")
- S57. (MH "Sleep-Wake Transition Disorders")
- S58. (MH "Stress, Occupational+")
- S59. (MH "Stress, Psychological")
- S60. ("24 hour\*" or "24 hr\*" or "twenty four hour\*" or "twentyfour hour\*") N1 rhythm\*
- S61. "biological rhythm\*"
- S62. "burn out\*" or "burned out\*" or "burnt out\*" or burnout\*
- S63. "circadian misalignment"
- S64. (circadian or diurnam or ultradian) N1 rhythm\*
- S65. exhaust\*
- S66. fatigu\*
- S67. sleep\* N3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)
- S68. tired\*
- S69. weariness
- S70. S49 OR S50 OR S51 OR S52 OR S53 OR S54 OR S55 OR S56 OR S57 OR S58 OR S59 OR S60 OR S61 OR
- S62 OR S63 OR S64 OR S65 OR S66 OR S67 OR S68 OR S69
- S71. S48 AND S70
- S72. S48 AND S70 Limiters Published Date: 20000101-20161231; Publication Type: Clinical Trial, Journal

Article, Meta Analysis, Meta Synthesis, Practice Guidelines, Randomized Controlled Trial, Research,

Review, Systematic Review; Language: English, French

Database: PubMed via NCBI Entrez

Date searched: 14 April 2016

**Records retrieved: 92** 

(((("Medical Staff, Hospital"[mh:noexp] OR "Physician Impairment"[mh:noexp] OR "Physicians"[mh] OR allergist[ti] OR allergists[ti] OR anaesthetist[ti] OR anaesthetists[ti] OR anaesthesiologist[ti] OR anaesthesiologists[ti] OR anesthetist[ti] OR anesthetists[ti] OR anesthesiologist[ti] OR anesthesiologists[ti] OR cardiologist[ti] OR cardiologists[ti] OR clinician[ti] OR clinicians[ti] OR ((clinician[tiab] OR clinicians[tiab]) AND (absent[tiab] OR absentee[tiab] OR absenteeism[tiab] OR absenteeisms[tiab] OR "burned out"[tiab] OR "burn out"[tiab] OR "burn outs"[tiab] OR burnout[tiab] OR burnouts[tiab] OR "burnt out"[tiab] OR coping[tiab] OR distress[tiab] OR distressed[tiab] OR distresses[tiab] OR distressing[tiab] OR "duty hour"[tiab] OR "duty hours"[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR health[tiab] OR healthy[tiab] OR impair[tiab] OR impaired[tiab] OR impairment[tiab] OR resilience[tiab] OR resiliency[tiab] OR resilient[tiab] OR satisfaction[tiab] OR sleep[tiab] OR sleepiness[tiab] OR sleeps[tiab] OR "well being"[tiab] OR wellbeing[tiab] OR wellness[tiab] OR "work hours"[tiab] OR "working hours"[tiab] OR "work life balance"[tiab])) OR ((clinician[tiab] OR clinicians[tiab]) AND (cities[tiab] OR city[tiab] OR community[tiab] OR communities[tiab] OR country[tiab] OR countryside[tiab] OR frontier[tiab] OR north[tiab] OR northern[tiab] OR remote[tiab] OR rural[tiab] OR suburb[tiab] OR suburbs[tiab] OR suburban[tiab] OR urban[tiab] OR urbanite[tiab])) OR dermatologist[ti] OR dermatologists[ti] OR endocrinologist[ti] OR endocrinologists[ti] OR doctor[ti] OR doctors[ti] OR ((doctor[tiab] OR doctors[tiab]) AND (absent[tiab] OR absentee[tiab] OR absenteeism[tiab] OR absenteeisms[tiab] OR "burned out"[tiab] OR "burn out"[tiab] OR "burn outs"[tiab] OR burnout[tiab] OR burnouts[tiab] OR "burnt out"[tiab] OR coping[tiab] OR distress[tiab] OR distressed[tiab] OR distresses[tiab] OR distressing[tiab] OR "duty hour"[tiab] OR "duty hours"[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR health[tiab] OR healthy[tiab] OR impair[tiab] OR impaired[tiab] OR impairment[tiab] OR resilience[tiab] OR resiliency[tiab] OR resilient[tiab] OR satisfaction[tiab] OR sleep[tiab] OR sleepiness[tiab] OR sleeps[tiab] OR "well being"[tiab] OR wellbeing[tiab] OR wellness[tiab] OR "work hours"[tiab] OR "working hours"[tiab] OR "work life balance"[tiab])) OR ((doctor[tiab] OR doctors[tiab]) AND (cities[tiab] OR city[tiab] OR community[tiab] OR communities[tiab] OR country[tiab] OR countryside[tiab] OR frontier[tiab] OR north[tiab] OR northern[tiab] OR remote[tiab] OR rural[tiab] OR suburb[tiab] OR suburbs[tiab] OR suburban[tiab] OR urban[tiab] OR urbanite[tiab])) OR "family practitioner"[ti] OR

"family practitioners"[ti] OR gastroenterologist[ti] OR gastroenterologists[ti] OR "general practice physician"[ti] OR "general practice physicians"[ti] OR "general practitioner"[ti] OR "general practitioners"[ti] OR geriatrician[ti] OR geriatricians[ti] OR gynaecologist[ti] OR gynaecologists[ti] OR gynecologist[ti] OR gynecologists[ti] OR haematologist[ti] OR haematologists[ti] OR hematologist[ti] OR hematologists[ti] OR "health care professional"[ti] OR "health care professionals"[ti] AND "health care provider"[ti] OR "health care providers" OR "health professional"[ti] OR "health professionals"[ti] OR "health provider"[ti] OR "health providers"[ti] OR "healthcare professional"[ti] OR "healthcare professionals"[ti] OR "healthcare provider"[ti] OR "healthcare providers"[ti] OR hospitalist[ti] OR hospitalists[ti] OR "house staff"[ti] OR "house staffs"[ti] OR housestaff[ti] OR housestaffs[ti] OR intensivist[ti] OR intensivists[ti] OR internist[ti] OR internists[ti] OR "medical professional"[ti] OR "medical professionals"[ti] OR obstetrician[ti] OR obstetricians[ti] OR oncologist[ti] OR oncologists[ti] OR ophthalmologist[ti] OR ophthalmologists[ti] OR orthopaedist[ti] OR orthopaedists[ti] OR orthopaedist[ti] OR orthopedists[ti] OR otolaryngologist[ti] OR otolaryngologists[ti] OR otorhinolaryngologist[ti] OR otorhinolaryngologists[ti] OR neonatologist[ti] OR neonatologists[ti] OR nephrologist[ti] OR nephrologists[ti] OR neurologist[ti] OR neurologists[ti] OR neuropsychiatrist[ti] OR neuropsychiatrists[ti] OR neurosurgeon[ti] OR neurosurgeons[ti] OR paediatrician[ti] OR paediatricians[ti] OR pediatrician[ti] OR pediatricians[ti] OR perinatologist[ti] OR perinatologists[ti] OR physicians[ti] OR physicians[ti] OR ((physician[tiab] OR physicians[tiab]) AND (absent[tiab] OR absentee[tiab] OR absenteeism[tiab] OR absenteeisms[tiab] OR "burned out"[tiab] OR "burn out"[tiab] OR "burn outs"[tiab] OR burnout[tiab] OR burnouts[tiab] OR "burnt out"[tiab] OR coping[tiab] OR distress[tiab] OR distressed[tiab] OR distresses[tiab] OR distressing[tiab] OR "duty hour"[tiab] OR "duty hours"[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR health[tiab] OR healthy[tiab] OR impair[tiab] OR impaired[tiab] OR impairment[tiab] OR resilience[tiab] OR resiliency[tiab] OR resilient[tiab] OR satisfaction[tiab] OR sleep[tiab] OR sleepiness[tiab] OR sleeps[tiab] OR "well being"[tiab] OR wellbeing[tiab] OR wellness[tiab] OR "work hours"[tiab] OR "working hours"[tiab] OR "work life balance"[tiab])) OR ((physician[tiab] OR physicians[tiab]) AND (cities[tiab] OR city[tiab] OR community[tiab] OR communities[tiab] OR country[tiab] OR countryside[tiab] OR frontier[tiab] OR north[tiab] OR northern[tiab] OR remote[tiab] OR rural[tiab] OR suburb[tiab] OR suburbs[tiab] OR suburban[tiab] OR urban[tiab] OR urbanite[tiab])) OR "primary care practitioner"[ti] OR "primary care practitioners"[ti] OR psychiatrist[ti] OR psychiatrists[ti] OR pulmonologist[ti] OR pulmonologists[ti] OR rheumatologist[ti] OR rheumatologists[ti] OR surgeon[ti] OR surgeons[ti] OR ((surgeon[tiab] OR surgeons[tiab]) AND (absent[tiab] OR absentee[tiab] OR absenteeism[tiab] OR absenteeisms[tiab] OR

"burned out"[tiab] OR "burn out"[tiab] OR "burn outs"[tiab] OR burnout[tiab] OR burnouts[tiab] OR

## **Supplementary table 1.** Descriptive characteristics of the included studies

				ВМЈ Ор	5/bmjopen-2018-021967 ccted by copyright, inducention			
Supplementary	table 1. Descriptive char	acterist	ics of the inc	luded studies			018-0219 yright, ir	
Study	Physician and patient chara	cteristics			Setting		Interventions or 3	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures on 21 Se	
Observational (exp	posure) studies (n=45)						epte Er	
Cohort design	, , ,						es r	
Chu, 2011 [32]	Surgeons	6	NR	Range: 32-55y	Tertiary care academic	Urban	Insufficient	Length of surgery; patient
Canada	Patients: cardiac surgery cases	4,047	NR	NR	- hospital		to work on 银色配件 preceding s破离子 ex	postoperative mortality, complications, length of stay
Ellman, 2004 [41] US	Surgeons Patients: adult cardiac surgery cases	NR 6,751	NR 70%	NR S: 63.4±0.7y C: 63.5±0.1y	University hospitals	Urban	Insufficient see Properties to work on the Bht preceding see Fro	Length of surgery; patient complications, in-hospital mortality, length of stay, need for blood products
Govindarajan,	Surgeons	1,448	NR	46.3±8.7	Academic and non-	Mixed	Sleep deprivation due	Length of surgery; Patient
2015 [31] Canada	Patients: surgical cases	38,978	NR	56.4±16.6y	academic hospitals	IVIIXEU	to work on the night preceding a Haytine surgery	complications, mortality, readmissions, length of stay
Rothschild, 2009 [58] US	Surgeons Obstetrician/gynecologists	220	Surgeons: 84% OB/GYNs: 28%	Surgeons: 42.0±7.6y OB/GYNs: 42.0±9.0y	Tertiary care academic trauma centre/referral centre for high-risk obstetrics	Urban	Sleep deprivation due to work on the night preceding a paytine	Patient complications, preventable complications
	Patients: surgical and obstetrics cases	Surg.: 4,471 Obst.: 4,902	Surg: S: 25% C: 28% Obst.: S: 0% C: 0%	Surg: S: 49.1±16.3y C: 50.0±16.3y Obst.: S: 32.9±5.2y C: 33.5±5.0y			d similar technologies	
Schieman, 2007 [63]	Colorectal surgeons	NR	NR	NR	University teaching hospitals	NR	Fatigue due to work	Length of surgery; patient operative complications,
Canada	Patients: undergoing anterior resection for rectal cancer	270	NR	S: 64.5y C: 64.4y	-		gence Bibliogr	length of stay, mortality, cancer recurrence
	For	peer rev	iew only - htt	p://bmjopen.b	mj.com/site/about/gu	idelines	Bibliographique de l	1

				вмл О	oen		6/bm jopen-2018-0219 cted by copyrightisin	
Study	Physician and patient chara	cteristics			Setting		Interventions or 21	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures including	
Vinden, 2014 [77] Canada	General surgeons Patients: Elective cholecystectomies	331 10,390	83% S: 27% C: 26%	48±10y S: 49±16y C: 49±16y	Community hospitals	Mixed	Sleep deprivation value to overnight was preceding day was surgery	Patient mortality, operative complications
<b>Before-after desigr</b> Amirian, 2014 [34] Denmark		29	55%	Median: 35y Range: 27-49y	Academic hospital	Urban	17-h night skifeweth sleep deprivation	Cognitive and psychomotor abilities on a laparoscopic simulatior
Gerdes, 2008 [45] US	Surgeons	9	NR	NR	University Hospital	Urban	Fatigue; sleep in a deprivation of the control of t	Cognitive and psychomotor abilities
Lederer, 2006 [50] Austria	Senior anesthetists	11	82%	49.0±2.0y	Hospital	Urban	Sleep deprive to specific from 24-h call ships to ships t	Concentration ability; reaction time; performance on psychometric tasks
Time series design							rain Jo	. ,
Leichtfried, 2011 [51] Austria	Anesthetists	10	100%	Mean: 32y Range: 29-35y	University Hospital	Urban	Sleep deprivations from 24-h slaft; sleepiness, Reepiness, hours	Melatonin metabolite profile
Cross-sectional des	sign						nii ar	
Aziz, 2004 [35] US	Family medicine physicians Various specialties	153	NR	NR	Hospitals	NR	Fatigue C 3	Stress
Beaujouan, 2005 [36] France	Anesthesiologists	3,476	64%	≤35y: 9% 36-45y: 28% 46-55y: 49% 56-65y: 13%	Public sector General hospitals University hospitals Private hospitals	NR	June 9, 2025 at Age	Substance abuse
Chang, 2013 [37] US	Anesthesiologists	11	64%	Mean: 38y IQR: 34-48y	Level 1 trauma centre	NR	Sleep deprivation due	Cognitive performance; reaction time
	For	peer rev	iew only - htt	p://bmjopen.b	mj.com/site/about/gu	iidelines	shift; sleepiness bill ographique a.xhtml	

			cted by copyrighting interventions					
Study	Physician and patient chara	cteristics			Setting	rights or 21	Outcomes	
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	967 on 2 including exposures	
Chen, 2008 [38] US	Psychiatrists Internists General practitioners Surgeons Obstetrician-gynecologists Radiologists Pediatricians Other	180	77%	Academic: 79% 36-55y Private practice: 73% 36-65y	Medical school Private practices	Urban	Sleep deprive uses related to text a	Impact on personal and professional life; perceived risk of errors
Doppia, 2011 [39] France	Anesthesiologists	565	64%	<35y: 11% 35-54y: 63% >55y: 25%	Public hospitals Private hospitals Work-health environments Public health units	NR	Sleep deprive data minin	Burnout
Elovaino, 2015 [42] Finland	Physicians in various specialties	1,524	40%	Median: 49.7y Range: 24-69y	Hospitals Primary care Private practice Other unspecified	NR	Sleep difficult training	Job demands and contro
Gander, 2000 [43] New Zealand	Anesthetists	183	NR	Mean: 46y	Combined public/private practice Other unspecified	NR	Work-related sleep disturbanced Si	Risk of fatigue-related errors
Harbeck, 2015 [46] Germany	Internists	20	45%	Median: 32y Range: 26-42y	Hospital	NR	Sleep disturtiance due to a 24-call shift on Jun	Biochemical and physiological parameters neurocognitive function
Heponiemi, 2014 [47] Finland	Physicians in various specialties Non-specialized physicians	1,541	40%	49.80±9.49y, Range: 24-67y	Hospitals Primary care clinic Private practice Other unspecified	NR	Sleep difficuties 🖁	Job satisfaction; work ability; psychological distress
	For	peer rev	iew only - htt:	o://bmionen.h	mj.com/site/about/gu	idelines	2025 at Agence Bibliographique de I gies.	

		5/bmjopen-2018-0219 cted by copyrightsin						
Study	Physician and patient chara	cteristics			Setting		Interventions or 22	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	including exposures	
Jackson, 2017 [48] US	Surgeons in various subspecialties	993	61%	More; less satisfied: 30-39y: 23%;24% 40-49y: 32%;36% 50-59y: 23%;27% ≥60y: 23%;14%	Academic practice Non-academic practice	NR	Steptember 2018. Downloade Enseignement Superieur War uses related to text and da	
Kanieta, 2011 [49] Japan	Internists Surgeons Orthopedics Pediatricians Obstetrician-gynecologists Psychiatrists Dermatologists Urologists Opthalmologists Otorhinolaryngologists Other	3,486	66%	20-39y: 11% 40-49y: 25% 50-59y: 28% 60-69y: 16% ≥70y: 21%	Hospitals Clinics Other unspecified	NR	Sleep deprite http://bmjopen.bmj.com/difficulties; Al training, and simil.	Medical incidents
Lindfors, 2006 [52] Finland	Anesthetists	328	53%	47±7.8y Range: 32-69y	University hospitals Central and district hospitals Private sector	NR	Sleep disturbances; sleepiness chnolog	Stress; suicidal tendencies
Mahmood, 2016 [53] Norway	Generalists Internists Pediatricians Surgical specialties Anesthesiologists	450 (all time points)	41%	43y±2.8y	Public health system Private practice	NR	Sleep deprivation at Agence Bibliographiq	e Alcohol misuse

		5/bmjopen-2018-0219 cted by copyrightsin						
Study	Physician and patient chara	acteristics			Setting	Interventions or 21	Outcomes	
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	967 on 2 including exposures	
Nishimura, 2014 [54] Japan	Neurosurgeons and neurologists	2,564	NR	NR	Stroke care centres Teaching hospitals	NR	Sleep deprior English	Burnout
Pit, 2014 [55] Australia	General practitioners	92	60%	50±10.7y	NR	Rural	Work-related and 20 disturbance of 20	Early retirement intentions
Pit, 2016 [56] Australia	General practitioners	92	60%	50±10.7y	Private (solo) practice Group practice	Rural	Work-relate <del>d</del> <b>⊈</b> e <b>⊊</b> p	Sickness presenteeism
Roberts, 2014 [57] US	General internists Internal medicine hospitalists	578	58%	Hospitalists: 46.9±12.4y Generalists: 53.6±10.2y	Private practice Academic medical centre Veterans hospital Military practice Other	NR	Downloaded from ht superieur (ABES). Fatigue	Falling asleep while driving
Saadat, 2016 [60] US	Anesthesiologists	21	71%	30-40y: 57% 41-50y: 19% 51-55y: 24% Range: 32-56y	Tertiary care academic children's hospital	NR	Sleep deprivation due to 17-h night-call hift aining a a	Mood disturbances
Saadat, 2017 [59] US	Anesthesiologists	21	65%	Range: 32-56 years	Tertiary care academic children's hospital	NR	Sleep deprivation due to 17-h night call shift	
Sanches, 2015 [61] Spain	Emergency medicine physicians	18	28%	29.2±2.6y	Central hospital	NR	Sleep deprivation on	Cognitive and psychomotor abilities
Sargent, 2009 [62] US	Orthopedic surgeons	264	92%	NR	Orthopedic surgery training programs	NR	Sleep deprivation 9, 20	Burnout; psychologica distress; marital satisfaction
Sende, 2012 [64] France	Emergency physicians	318	62%	39±8y	Hospitals Mobile emergency services Other unspecified	NR	Fatigue; sleep 25 at Agence Bibliographique de l	Stress

				ВМЈО	pen		Intervention	/bmjopen-201	
Study	Physician and patient cl	naracteristics		Setting				8-021g	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures of	967 on 21	
Sexton, 2001 [65] US	Consulting physicians: Surgeons Anesthesiologists Pulmonary physicians Cardiologists Pediatricians	271	NR	NR	Teaching and non- teaching hospitals	Urban	Fatigue Or uses related	ptember 201 Enseignem	Perceived performance effectiveness
Shanafelt, 2005 [67] US, Canada, Mexico	Oncologists	241	85%	>50y: 51%	Community clinics Hospitals Private practice Academic medical centres	NR	Fatigue; sleet deprivation	Downloaded fi	Quality of life/well-being
Shanafelt, 2010 [66] US	Surgeons	7,905	87%	Median: 51y Q1: 43y Q2: 59y	Private practice Academic medical centres Veterans hospital Active military practice Retired or not in practice Other	NR	Fatigue  Fatigue  Fatigue  Fatigue  Fatigue  Fatigue	nom http://bmjopen.b	Perceived major medical errors
Shanafelt, 2014 [68] US	Oncologists	1,117	52%	Median: 52y	Private practice Academic practice Veteran's hospital Industry, other	NR	Fatigue of	mj.com/ on	Satisfaction with work-lif balance
Shirom, 2006 [69] Israel	Opthalmologists Dermatologists Otolaryngologists Gynecologists General surgeons Cardiologists	890	80%	Median: 52y SD: 7.2y	Community clinics Acute care hospital outpatient clinics	NR	Physical fatig	June 9, 2025 at Agence Bibliographique de l	Perception of quality of patient care
		For peer rev	iew only - htt <sub>l</sub>	o://bmjopen.k	omj.com/site/about/gu	idelines.	xhtml	)graphique de l	

				6/bmjopen-2018- cted by copyrigl				
Study	Physician and patient char	acteristics			Setting		Interventions or 2	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	including exposures	
Shirom, 2010 [70] Israel	Opthalmologists Dermatologists Otolaryngologists Gynecologists General surgeons Cardiologists	890	80%	Median: 52y SD: 7.2y	Community clinics Acute care hospital outpatient clinics	NR	Physical fatigue Enseignemer 2018.	Burnout
Smith, 2017 [71] UK	General practitioners Surgeons Other unspecified specialties	3,550	63%	NR	NR (varied)	NR	Perceived fate Down sleep deprivation da	Physical and mental health; competence
Starmer, 2016 [72] US	General pediatricians Pediatric surgeons Pediatric hospitalists Pediatric specialists (unspecified)	840	40%	NR	NR (some in private practice)	NR	Sleep deprive mining,	Burnout; balanced personal and professional commitments; life and career satisfaction
Tanti, 2017 [73] Malta	Physicians (unspecified)	204	62%	Median: 41y	Hospitals Community Office-based	NR	Fatigue Fatigue a significant parameter	Prescribing errors
Tokuda, 2009 [74] Japan	Hospital physicians: Generalists Other unspecified specialties	236	75%	40.9±7.8y Range: 26-76y	Hospitals with ≥20 inpatient beds	NR	Sleep deprivation com/ on a	Burnout; job satisfaction
Vela-Bueno, 2008 [76] Spain	Primary care physicians	113	27%	41.4±8.0y	Primary care centres	Urban	Sleep problems, une 9, 20	Burnout
Wada, 2010 [78] Japan	Physicians (unspecified)	3,862	78%	M: 75% 30- 59y F: 85% 30-59y	Hospitals	NR	025 at Agence Bibliographic	Depressive symptoms
							araphi	7

				ВМЈ О	pen		6/bmjopen-2018 cted by copyrig	
Study	Physician and patient cha	ractoristics			Setting		Interventions or 2	Outcomes
Country	Type	n=	Sex (% male)	Δσρ	Location	Urban	- 5 0	Outcomes
.ountry	Type	.,-	Sex (70 maie)	786	Location	or		
						rural	on 2	
lon-comparative o	design						<u> </u>	
ander, 2008 [43]		20	85%	Median: 44y	Hospitals	Urban	Sleep distur and	Psychomotor
ew Zealand				,			from consecutive	performance
							working day	p
							call work te	
tervention studie	es (n=2)						# # # # # # # # # # # # # # # # # # #	
andomized contro							o te	
	Emergency physicians	17	35%	39.1y±6.9y	University hospital	Urban		Perceived stress; urine
ance	- 0,			- ,,	/ · · P · · · ·		and 24-h sh	interleukine-8
							deprivation 200 deep	
							quality;	
chal, 2005 [75]	Surgeons	64	67%	Median:	Government hospitals	NR	Sleep deprive	Product quality,
orway	Gynecologists	04	0770	Post-call:	Government nospituis	1411	to 24-h call ghift	procedure effectiveness
orway	Orthopedic surgeons			33.0y			7 0	of a surgical simulation
	Urologists			Post-work:			A 🕌	or a surfical simulation
	Vascular surgeons			38.0y			bmjo train	
nitea Kingaom; US	5: United States of America	; y: year(s)					bmj.com/ on June 9, 2025 at Agence and similar technologies.	
	F	or peer rev	iew only - htt	p://bmjopen.b	omj.com/site/about/gu	idelines.	2025 at Agence Bibliographique de I gies.	

### Supplementary file 3. Risk of bias assessments

Summary of risk of bias assessments for randomized controlled trials (n=2)<sup>a</sup>

First Author, Year	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other sources of bias	Overall risk of bias <sup>b</sup>
Dutheil, 2013	Low	Unclear	High	High	Low	Low	High	High
Uchal, 2005	Low	Low	Unclear	Low	Low	Low	Low	Unclear

<sup>&</sup>lt;sup>a</sup>Assessed using the Cochrane Collaboration's Risk of Bias Tool

Summary of quality assessments for cohort studies (n=6)<sup>a</sup>

First Author,	Selection					Comparab	oility		Outco	me		Total
Year	Representa-	Selection	Ascertain-	Outcome	Total	Compara	Total	Assess-	Adequate	Adequate	Total	Scoreb
	tiveness of	of non-	ment of	not	/4	-bility of	/2	ment of	length of	follow-up	/1	/9
	exposed	exposed	exposure	present at		cohorts		outcome	follow-up	of cohorts		
	cohort	cohort	/1	start		/2		/1	/1	/1		
	/1	/1		/1								
Chu, 2011	1	1	0	1	3	2	2	1	1	1	3	8
Ellman, 2004	1	1	1	1	4	1	1	1	1	1	3	8
Govindarajan,	1	1	1	1	4	2	2	1	1	1	3	9
2015												
Rothschild, 2009	1	1	1	1	4	2	2	1	1	1	3	9
Schieman, 2008	1	1	1	1	4	1	1	1	1	1	3	8
Vinden, 2014	1	1	1	1	4	1	1	1	1	1	3	8

<sup>&</sup>lt;sup>a</sup>Assessed using the Newcastle-Ottawa Quality Assessment Scale

<sup>&</sup>lt;sup>b</sup>Overall risk of bias is Low if all domains are rated as low, High if at least one domain is assessed as high, and Unclear if at least one domain is assessed as unclear and no domains are assessed as high

<sup>&</sup>lt;sup>b</sup>An overall score of 7 to 9 stars is considered as low risk of bias, 4 to 6 as unclear risk of bias, and 3 or less as high risk of bias

Summary of risk of bias assessments for before-after studies (n=3)<sup>a</sup>

First Author, Year	Random sequence generation <sup>b</sup>	Allocation concealment <sup>b</sup>	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other sources of bias <sup>c</sup>	Overall risk of bias <sup>d</sup>
Amirian, 2014	NA	NA	High	High	Low	Low	High	High
Gerdes, 2008	NA	NA	High	High	Low	Low	High	High
Lederer, 2006	NA	NA	High	High	Low	Low	High	High

<sup>&</sup>lt;sup>a</sup>Assessed using Cochrane Effective Practice and Organization of Care (EPOC) Review Group's criteria for before-after studies, adapted from the Cochrane Collaboration Risk of Bias Tool

# Summary of risk of bias assessments for time series studies (n=1)<sup>a</sup>

First Author, Year	Intervention independent of other changes	Intervention effect pre- specified	Intervention unlikely to affect data collection	Allocation concealment <sup>a</sup>	Incomplete outcome data	Selective reporting	Other sources of bias <sup>c</sup>	Overall risk of bias <sup>d</sup>
Leitchfried, 2011	Low	High	Low	NA	Low	Low	High	High

<sup>&</sup>lt;sup>a</sup>Assessed using Cochrane Effective Practice and Organization of Care (EPOC) Review Group's criteria for interrupted time series studies, adapted from the Cochrane Collaboration Risk of Bias Tool

<sup>&</sup>lt;sup>b</sup>Assessed as 'not applicable' (NA) when the studies did not include a control group

<sup>&</sup>lt;sup>c</sup>Assessed as High due to lack of a control group

<sup>&</sup>lt;sup>d</sup>Overall risk of bias is Low if all domains are rated as low, High if at least one domain is assessed as high, and Unclear if at least one domain is assessed as unclear and no domains are assessed as high

<sup>&</sup>lt;sup>b</sup>Assessed as not applicable (NA) when the studies did not include a control group

<sup>&</sup>lt;sup>c</sup>Assessed as High due to lack of a control group

<sup>&</sup>lt;sup>d</sup>Overall risk of bias is Low if all domains are rated as low, High if at least one domain is assessed as high, and Unclear if at least one domain is assessed as unclear and no domains are assessed as high

First Author, Year		Selection			Outcome			Total Score <sup>b</sup>
	Adequacy of case definition /1	Representative- ness of the sample /1	Total /2	Assessment of outcome /1	Same method of ascertainment for entire sample /1	Response rate /1	Total /3	/5
Aziz, 2004	0	0	0	0	1	0	1	1
Beaujouan, 2005	1	0	1	0	1	0	1	2
Chang, 2013	1	0	1	0	1	1	2	3
Chen, 2008	1	0	1	0	1	0	1	2
Doppia, 2011	1	1	2	0	1	1	2	4
Elovaino, 2015	1	1	2	0	1	1	2	4
Gander, 2000	1	1	2	0	1	1	2	4
Harbeck, 2015	1	0	1	0	1	1	2	3
Heponiemi, 2014	1	1	2	0	1	1	2	4
Jackson, 2017	0	0	0	0	1	0	1	1
Kanieta, 2011	1	0	1	0	1	1	2	3
Lindfors, 2006	1	1	2	0		1	2	4
Mahmood, 2017	1	0	1	0	1	0	1	2
Nishimura, 2014	1	1	2	0	1	0	1	3
Pit, 2014	1	0	1	0	1	1	2	3
Pit, 2016	1	0	1	0	1	1	2	3
Roberts, 2014	1	1	2	0	1	0	1	3
Saadat, 2016	1	1	2	0	1	1	2	4
Saadat, 2017	1	1	2	0	1	1	2	4
Sanches, 2015	1	0	1	0	1	0	1	2
Sargent, 2009	1	0	1	0	1	0	1	2

First Author, Year		Selection			Outcome			Total Score <sup>b</sup>
	Adequacy of	Representative-	Total	Assessment of	Same method of	Response rate	Total	/5
	case definition	ness of the sample	/2	outcome	ascertainment for	/1	/3	
	/1	/1		/1	entire sample			
					/1			
Sende, 2010	1	0	1	0	1	0	1	2
Sexton, 2001	1	0	1	0	1	0	1	2
Shanafelt, 2005	1	0	1	0	1	1	2	3
Shanafelt, 2010	1	1	2	0	1	0	1	3
Shanafelt, 2014	1	0	1	0	1	1	2	3
Shirom, 2006	1	1	2	0	1	1	2	4
Shirom, 2010	1	1	2	0	1	1	2	4
Smith, 2016	1	0	1	0	1	1	2	3
Starmer, 2016	1	1	2	0	1	1	2	4
Tanti, 2017	1	0	1	0	1	0	1	2
Tokuda, 2009	1	1	2	0	1	1	2	4
Vela-Bueno, 2008	1	1	2	0	1	1	2	4
Wada, 2010	1	1	2	0	1	0	1	3

<sup>&</sup>lt;sup>a</sup>Assessed using the Newcastle-Ottawa Quality Assessment Scale, adapted for cross-sectional studies

<sup>&</sup>lt;sup>b</sup>An overall score of 4 to 5 stars is considered as low risk of bias, 3 as unclear risk of bias, and 2 or less as high risk of bias. For response rate, ≥50% was used as the criterion to be awarded a star

Summary of quality assessments for non-comparative studies (n=1)<sup>a</sup>

First Author,		Selection		Exposu	re		Outcome			Total
Year	Adequacy of case definition /1	Representat- iveness of the sample /1	Total /2	Ascertain- ment of exposure	Total /1	Assessment of outcome /1	Same method of assessment for entire sample	Loss to follow-up /1	Total /3	Score <sup>b</sup> /6
Gander, 2008	1	1	2	0	0	0	<b>/1</b> 1	1	2	4

<sup>&</sup>lt;sup>a</sup>Assessed using the Newcastle-Ottawa Quality Assessment Scale, adapted by the authors to be suitable to the non-comparative design

<sup>&</sup>lt;sup>b</sup>An overall score of 5 to 6 stars is considered as low risk of bias, 3 to 4 as unclear risk of bias, and 2 or less as high risk of bias 

## Supplementary file 4. Detailed study outcomes

### Physician health and wellness outcomes and associations with fatigue

			ВМЈ Оре	n	3/bmjopen-2018-021967 cted by copyright, inclu
		etailed study outcomes	associations with fatigue		5/bmjopen-2018-021967 on 2 cted by copyright, including
Study	Study	Exposures or interventions		Outcomes	<u>യ</u> <u>ഉ</u> Associations getween exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	<del>_</del>
		time points		time points	ptem Ensuuses
Surgeons					rei:
Jackson, 2017	CS	Not feeling well rested: self-	71% healthy, 28% unhealthy in	Job satisfaction: Abridged	Job satisfaction hose more vs. less satisfied:
		reported as 'unhealthy'	terms of being well rested	Job in General Scale;	Healthy (wel हैं क्वें क्वें): 85% vs. 58%, p<0001;
RoB: high		Time points NR		grouped into more or less satisfied using the median	Unhealthy (nat វ៉ាម្ហា rested): 15% vs. 42%, p<0.00:
				Time a mainte ND	oad nd c
Nishimura, 2014	CS	Sleep hours/night: self-	Mean±SD sleep: 5.94±1.08h	Time points NR Burnout: Japanese MBI	1) Mean±SD leep±or not burned out vs. mild to
Mishimura, 2014	CS	reported (continuous)	Wednesd sleep. 5.94±1.0611	(severe: EE >4.0 and either	moderate vs 3 (2002) moderate
RoB: unclear		reported (continuous)		DP >2.6 or PE <4.17)	5.63±0.94, p. 20.05
		Time points NR			2) Association between sleep and burnout (OR
		•		Time points NR	(95% CI)): bivariato 0.67 (0.61-0.73), p<0.001;
					multivariate 🖺 clu🖥 ng work characteristics and
					mental healt (0.75-0.94), p=0.002.
Sargent, 2009	CS	Sleep deprivation: self-	21% none, 48% a little, 23%	Burnout: MBI (norms NR);	1) Positive correlation between sleep deprivation
		reported on a 4-point scale	quite a bit, 8% a lot	Marital satisfaction: RDAS;	and EE, DP, paychological distress, lower marital
RoB: high		(none, a little, quite a bit, a		Psychological morbidity:	satisfaction, gi p 2.001. No relationship with PA.
		lot)		GHQ-12 score ≥4	/ or
		Time points NR		Time points NR	tech
Anesthesiologists <sup>a</sup>		Time points NK		Time points WK	ch une
Lederer, 2006	BA	24-h shift with on-call duty;	Mean±SD sleep: 4.1±1.7h;	Stress during duty: 4-point	1) Mean stress score during duty: 2.1.
2000	<i>57</i> (	Sleep hours and	Number of interruptions:	scale from 'calm' to 'very	s S
RoB: high		interruptions: self-reported;	0.8±1.1;	demanding'	
		Tiredness: VAS from 0 (low)	Tiredness pre- vs. post-duty:		Ag
		to 100 (high)	30.9±27.5 vs. 59.5±18.9,	Assessed post-duty	enc
			p=0.01.		ě m
		Assessed pre- and post-duty			Sib Di
		For peer revi	iew only - http://bmjopen.bm	j.com/site/about/guideline	at Agence Bibliographique de l

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and time points	Baseline	Assessment measure and time points	967 or
Leitchtfried, 2011	TS	24-h shift; Sleepiness: ESS (range: 0-	ESS (mean (range)): 7.4 (4-12); Mean±SD sleep hours:	aMT6-s: urinalysis	1) aMT6-s over shift, mean (95% CI): higher at 11:00AM pre (12\alpha (6.3-8.1)) and post-shift (9.3
RoB: high		24);	1) pre-study: 7.74±1.35h;	Assessed at 4-h intervals	(3.7-14.9)) v <b>⊊</b> d <b>;</b> q <b>;</b> p=0.016;
		Sleep hours: self-reported	2) Pre-24-h shift (11h00 on day	from 07:00 to 11:00	2) Correlation ween sleep and aMT6-s (data
		(continuous)	1: 0.13±0.35h, 19:00 on day 1:		NR): mild for the night prior with
			6.99±0.68h);		aMT6-s at 3PM he following day; sleep on night 2
		Sleepiness assessed pre-shit,	3) During the 24-h shift (07h00		with aMT6-s at a the next day; total sleep with
		sleep hours pre, during and post-shift	on day 2: 0.0±0.0h, 19h00 on		aMT6-s at 11♣ M So third day; moderate for sleep on first night → Som T6-s at 7AM and 11AM pre-
		post-stillt	day 2, 5.49±1.95h); 4) Post-24-h shift (11h00 on		shift, 11PM @ 24-h shift and 11AM post-shift
			day 3: 0.5±0.71h, 19h00 on		total sleep pe মান্ত and nocturnal sleep during 24
			day 3: 7.06±1.18h).		h shift with a at 11PM during shift; total
					sleep with a 🔁 🙀 🚉 3PM on first and second day
					11PM on section day;
					3) Correlations between ESS and aMT6-s:
				11	moderate fo (a) M (b) -s at 7AM during shift, 11AM on day off.
Beaujouan, 2005	CS	Sleep deprivation: 4-point	48.8% always or frequently	Substance abuse: 93-item	1) 60.6% with dependence vs. 46.0% of thos
DaD. biah		scale (always, frequently,	feel sleep deprived	addiction and substance	without reported deep difficulties, p<0.001.
RoB: high		rarely, never)		abuse questionnaire	2) OR (95% Ca of addiction for frequently/always vs. rarely/ne\( r \) sleep deprived: tobacco 1.42
		Time points NR		Time points NR	(1.04-1.94); (2.12-5.02).
Doppia, 2011	CS	Insufficient sleep: 4-point	28.9% reported insufficient	Burnout: CBI (mild: 1-2.4,	1) Frequency of bernout by response for sleep
		scale (no, not really, sort of,	sleep during work time	moderate: 2.5-3.5, severe:	sufficiency: 48.6% or no/not really, 16.3% for sor
RoB: low		yes)		3.6-5)	of/yes, p<0.@1. 👸
		Time points NR		Time points NR	)25 at
Lindfors, 2006	CS	Sleep hours/day: self-	Sleep hours (mean (range)): 7	Stress: MOSQ on a 3-point	1) Sleep sufficien predicted stress symptoms:
		reported to the nearest	(5-9)	scale (no, to some extent,	bivariate β=-0.36 <b>2</b> p<0.001; multivariate includin
RoB: low		0.5h;		clearly); Thoughts of suicide: 4-point	gender, sick leave suicide β=-0.269, p<0.001; 2) Sleep disturbange associated with thoughts of
		Adequacy of sleep and rest: self-reported (yes/no)		scale ('never' to 'have tried')	<u>~</u>

	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and time points	Baseline	Assessment measure and time points	967 on
		Time points NR		Time points NR	ng for
Saadat, 2015	CS	Sleep deprivation (<7h/24-h) due to 17-h overnight shift;	Mean±SD sleepiness on a regular day vs. post-call day:	Simple cognitive tests: VAS from 0 (not at all) to 100	Regular day र्ह्न भूजिन्द्र call day, mean±SD scores: 1) Simple cognitive tests: energetic 6.04±2.27 vs
RoB: low		Sleepiness and alertness: VAS from 0 (not at all) to 100 (extremely)	2.99±2.18 vs. 6.79±2.30, p<0.001	(extremely); Mood disturbance: PMS (scoring NR)	2.53±1.87, com and to 7.03±1.83 vs. 4.98±2.29, irritable 2.03 (2.99±2.16, sleepy 2.99±2. vs. 6.79±2.30 (2.99±2.174 vs. 2.41±1.97, p<0.001; jittery 7.7 (2.90±1.74 vs. 3.12±2.34, p=0.00
		All assessed on a regular day and a post-call day		All assessed on a regular day and a post-call day	anxiousness
					p<0.001; consistent 14±2.65 vs. 20.05±6.67, p=0.025; vige 24.05±6.75 vs.16.67±5.70, p<0.00 depression: hs; total mood disturbance: 42.57±15.26 vs. 76.90±6.91, p<0.001.
ER or ICU physiciar	ıs				nain jo
Dutheil, 2013 RoB: high	RCT	14-h or 24-h shift; Sleep hours: self-reported sleep and wake time;	1) Sleep duration and quality lower during shifts (14h and 24h) than any other day, and	Stress: VAS from 0 (low) to 100 (high); IL-8: urinalysis	1) Stress: higher following 14-h and 24-h shifts with the control day, ps0.05 (data NR); 2) IL-8: higher following 24-h shift vs. control
		Sleep quality: VAS from 1 (low) to 100 (high); Mental and physical fatigue:	lower during the 24-h vs. 14-h shift (p<0.05); 2) Mental and physical fatigue	Assessed at 08:30 and 18:30 on each day of protocol	(p=0.007) an 14-15 shift (p=0.015); ns difference between 14-15 shift and control day; 3) Correlations with IL-8: sleep hours pre-24-h
		VAS from 1 (low) to 100 (high)	higher after 14-h and 24-h shift vs. control day (data NR).	on each day of protocol	shift, r=-0.62, p=9.007; poor sleep quality during 14-h and 24-b shifts, r=0.452, p=0.031; 4) Multivaria le regression: 24-h shift increased
		Assessed on day prior to shift; during shift; each day of protocol (work, off, clerical, control)			8 by 1.9ng vs. con soliday, p=0.007; ns associati with 14-h shift, mental or physical fatigue, sleep deprivation, 14-h hit.
	CS	Fatigue and sleep	NR	Most important sources of	1) 78% indicated that sleep loss and fatigue were sources of stress.

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	
` '		time points		time points	udi
				related, organizational,	67 on 21 Se
		Time points NR		individual)	or u
				Time points NR	September Enseigi for uses rela
Generalists <sup>b</sup>					r 20 Jatec
Harbeck, 2015	CS	24-hours on-call shift with	1) Sleep hours on a normal day	Biochemical (laboratory	Before a normal shift vs. after overnight call shift:
		sleep disturbance: self-	vs. following a 24-h shift:	values) and physiological	1) Biochemica parameters: no changes in any
RoB: unclear		reported number of sleep	<2 hours: 0 vs. 5.9%; 2-4	(heart rate variability, skin	parameter eर्क्क्रिक्क्वाईor thyroid stimulating hormon
		disturbances and hours of	hours: 5.9% vs. 47.1%; 4-6	resistance, blood pressure)	which was high $\frac{1}{8}$ fter the on-call shift (p = 0.049)
		sleep per night	hours: 11.8% vs. 35.3%; >6	stress parameters	data NR); data (data NR);
			hours: 82.4% vs. 11.8%		data NR);
		Assessed before a normal	2) Number of sleep	Assessed before a normal	in any paramater a
		day shift, and after a 24-h on call shift	disturbances a normal day vs.	day shift, and after a 24-h on call shift	. <del>I</del>
		Call Stillt	following a 24-h shift: 0: 82.4% vs. 11.8%; 1: 11.8%	Off Call Still C	Al t
			vs. 35.3%; 2: 5.9% vs. 47.1%; 3:		raii 💆
			0% vs. 5.9%; 4: 0% vs. 0%; >4:		http://bmjopen ) . ng, Al training,
			0% vs. 0%		g, n. a b
Pit, 2014	CS	Work-related sleep	Work-related sleep	Early retirement (<65 years)	For sleep distarbance a few times a year to every
		disturbance: 7-point scale	disturbance: 41% never, 59% a	intentions (yes/no)	day vs. never
RoB: unclear		from 'never' to 'every day'	few times a year to every day		1) Intention के retire early: 74% vs. 26%, p<0.01;
		T' ' ' ND		Time points NR	2) Association with intention to retire early (OR
		Time points NR			(95% CI)): ungariase 3.6 (1.47-8.80), p<0.01;
					multivariate Acluumg work, occupational, individual facors 891 (1.11-7.6), p<0.05;
					4) RR (95% C) for tention to retire early: 2.0
					(1.18-3.49); attrib≌table fraction: 50.0%;
					population attributable fraction: 37.1%.
Pit, 2016	CS	Work-related sleep	Work-related sleep	Sickness presenteeism: 'yes'	For sleep disturbate a few times a year to every
		disturbance: 7-point scale	disturbance: 41% never, 59% a	response indicated 1 or	day vs. never:
RoB: unclear		from 'never' to 'every day'	few times a year to every day	more days	1) Sickness preser deeism: 32% vs. 68%, p=0.018;

cted by copyrigh 6/bmjopen-2018-0

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	- 967
		time points		time points	ludi
		Time points NR		Assessed for the past 12	2) Association with sickness presenteeism (OR
				months	(95% CI)): 2.92 (1.49-7.16), p=0.02.
Roberts, 2014	CS	Fatigue: LAS from 0 (low) to	Mean (SD) score: 5.8 (2.4) for	Impact of fatigue on daily	1) 8.7% of hospitalists and 4.3% of outpatient
		10 (high)	hospitalists; 5.9 (2.4) for	activities (falling asleep	general interdiscs and fallen asleep while driving
RoB: unclear			general internists	while driving) (yes/no)	due to fatigue
		Assessed for the past week			20 atec
				Time points NR	18.
Vela-Bueno, 2008	CS	Sleep Quality: PSQI	Prevalence (% (95% CI)):	Burnout: PBM with a 7-point	Low vs. high குயிர் <b>த</b> ut, mean±SD:
		(Spanish): score ≥5 indicates	1) Sleep-onset latency >30	scale from 1 (never) to 7	1) Global PSO = 5 ±2.22 vs. 7.24±4.17, p<0.001;
RoB: low		low quality (range; 0 to 21);	minutes: 8.4 (4.8-11.9);	(always)	2) PSQI subs (2) rescussion of the contract of
		Insomnia: DSM-IV criteria	2) Wake time after sleep onset		1.40±0.83, p <b>្នា</b> ១៤៩, sleep latency: 0.51±0.80 vs.
			>30 minutes: 15.4 (10.8-19.9);	Time points NR	$1.38\pm1.03$ , p $302$ ; sleep duration: $0.45\pm0.64$ vs.
		Time points NR; insomnia	3) Early morning awakening:		1.16±0.92, p 3 km s; sleep efficiency: 0.21±0.57 vs.
		symptoms in past month	22.5 (19.5-30.4);		0.77±0.98, p=0.012; sleep disturbance: ns; use of
			4) Nonrestorative sleep: 22.5		medication: 0.14 0.49 vs. 0.57 ±0.83, p=0.032;
			(17.2-27.7);		daytime dysfunction: 0.52±0.73 vs. 1.57±0.88,
			5) Daytime impairment for ≥5		p=0.002. <b>a</b> . <b>b</b> .
			days in past month: 14.2 (9.7-		3) Prevalenc (95 CI) of insomnia symptoms:
			18.6);		sleep latency 5.5 (2.5-11.5%) vs. 21.1% (10.5-
			6) Insomnia: 18.8 (13.8-23.7).		31.6%), p=0. 🗖 5; 👼 ake time > 30 min after sleep
					onset: 9.4% (∰:6-1 <mark>2</mark> .1%) vs. 25.5% (14.2-37.7%),
					p=0.029; ear awakening: 14.5% (5.1-23.8%) vs.
					45.6 (32.7-58 4%) p<0.001; somewhat/very
					dissatisfied with seep: 5.5% (2.5-11.5%) vs. 50%
					(37.1-62.8%) o <0 0 0 1; day impairment: 5.5% (2.5-
					11.5%) vs. 3 2 % (25.6-50.7%), p<0.001; insomnia:
					7.3% (0.4-14 <b>%</b> ) vs <b>2</b> 9.7% (27.1-52.2%), p<0.001.
Oncologists					at
Shanafelt, 2005	CS	Fatigue: LASA QOL ≤7;	75% had a high level of fatigue;	Wellbeing: 10-item LASA	1) Sleep deprivatian for high vs. low overall well-
		Sleep deprivation: 10-point	Mean±SD sleep score:	QOL, high ≥8 vs. low ≤7	being (mean±SD): 3.9±2.57 vs. 5.1±2.60, p=0.0004
RoB: unclear		Likert scale from 0 (not at	4.5±2.65.		2) Lower fatigue pedicted overall wellbeing in a
		all) to 10 (stressful as can		Time points NR	multivariate mod including personal and
		be)			professional characteristics, p=0.002.
					graphique de
					ihiq
					<b>ue</b>
		For peer revi	ew only - http://bmjopen.bmj	.com/site/about/guidelines	s.xntmi <u>e</u>

Study	Study	Exposures or interventions		Outcomes	cted by copyrigh 8-0  Associations 15 etween exposure and outcome
Risk of Bias (RoB)	•	Assessment measure and	Baseline	Assessment measure and	
0. 2.00 (2)		time points		time points	)67 or
					ing :
		Time points NR			ror φ
Shanafelt, 2014	CS	Fatigue: 10-point LAS (lower	Mean±SD fatigue score:	Satisfaction with WLB: 5-	1) OR (95%C for Satisfaction predicted by
		scores indicate greater	5.7±2.4	point Likert scale from	high fatigue (အန္တာကမ္ဘီး) in multivariate model
RoB: unclear		fatigue)		'strongly agree' to 'strongly	including person amand work-related factors, and
				disagree'	burnout: 0.4 කින් ජී ප්රථ (10), p<0.001.
		Time points NR		The area into ND	to to
Mixed groups of ph	veiciane			Time points NR	(e) SO
Aziz, 2004	CS	Working while fatigued: 5-	NR	Stress: 47-item	1) Sources of the grant working while fatigued had
1212, 2004	CJ	point scale from 'extreme'		questionnaire with a 5-point	mean±SD scoce at the score of
RoB: high		to 'a little'		scale from 'extreme' to 'a	0.653, in factor analysis;
. 0				little'	2) Inverse contains on between stress and working
		Time points NR			while fatigues: 1=2.270 (significance level NR).
				Time points NR	9, <b>A</b>
Chen, 2008	CS	Sleepiness: ESS score ≥11	Mean±SD ESS score: 7.8±4.0,	Impact on work and	1) Impact score correlated with ESS, r=0.31,
			range: 0-20, 23% had scores	personal life: Impact	p<0.05; <b>a</b> <del>o</del>
RoB: high		Time points NR	≥11.	Questionnaire with a 5-	2) ESS score an among physicians who
				point Likert scale from 1	agree/strongly agree vs. other response: worried
				(strongly agree) to 5	about having caraccident while driving home
				(strongly disagree)	post-call: 5.4 s. 7 p<0.001; sleep loss has a
					major impacano personal life: 8.4 vs. 7.0, p=0.01
				Time points NR	3) Higher ESS cores predicted by impact score in
					multivariate Beression including personal and
Elovaino, 2015	CS	Sleeping problems: Jenkins	Mean±SD score:	Jobs demands: 5 items	work-related dact gs: β=0.11, p=0.005.  There was n cassociation between sleeping
_10Va1110, 2013	CS	Scale with a 6-point scale	2006: 2.30 (1.00);	scored on a 5-point scale	problems in 2006 and job demands or control in
RoB: low		from 1 (never) to 6 (every	2010: 2.35 (1.05).	from 1 (strongly disagree) to	2010. <u>a</u>
		night)	2010: 2:00 (2:00):	5 (strongly agree);	:-
		64		Job control: 3 items derived	gen
		Assessed in 2006 and 2010		from the Karasek Job	Če
				Questionnaire	Bib
					Agence Bibliographique de l

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	967 ncli
		time points		time points	udi:
Heponiemi, 2014	CS	Sleeping problems: Jenkins	Mean±SD (range) score:	Psychological distress: GHQ-	1) Sleeping problems associated with job
		Scale <sup>81</sup> with a 6-point scale	2.30±1.00 (1-6)	12 with a 4-point scale (low	satisfaction, 📴 -0.🕰, p<0.001, psychological
RoB: low		from 1 (never) to 6 (every		to high);	distress, β=0 <b>5</b> ,8π <b>2</b> 0.001;
		night)		Job satisfaction: JDS with a	2) Total indir creffect of on-call duty through two
		1: 2006		Likert scale from 1 (strongly	mediators (see problems, work interference
		Assessed in 2006		disagree) to 5 (strongly	with family) ( CI)): job satisfaction 0.06 (-
				agree)	0.059, -0.016 (a) (a) (b) (c) (c) (c) (c) (c) (c) (c) (c) (c) (c
				Assessed in 2010	(0.023, 0.081) (0.001.
Mahmood, 2016	CS	Sleep deprivation: self-	Mean±SD hours:	Alcohol use disorders:	<u>্র ত ১</u> There was nন্দ্রহুক্ত্রiation between hours of sleep
Maiiiioou, 2010	C3	reported mean hours of	4 years: 4.52 (2.79);	Modified 9-item version of	when on call and parardous drinking behaviours
RoB: high		sleep when on call	10 years: 5.38 (6.36);	the Alcohol Use Disorder	(p=0.732) <b>a p</b>
		ordep when on dan	15 years: 6.41 (7.14).	Identification Test (AUDIT)	"
		Assessed at 4 years, 10		≥6 for men and ≥5 for	nin (S)
		years, and 15 years post-		women.	6, 7, <del>tp</del> :
		graduation			om http://bmjopen BES) . mining, Al training.
				Assessed at 4 years, 10	ain <mark>nj</mark> o
				years, and 15 years post-	ing per
				graduation	<u> </u>
Shirom, 2010	CS	Tiredness and exhaustion:	NR	Burnout: SMBM on a 7-	1) Correlation between physical fatigue subscale
		SMBM Physician Fatigue		point scale from 1 (almost	and overall but 10.88, p<0.05;
RoB: low		Subscale on a 7-point scale		never) to 7 (always)	2) In a predictive structural model for burnout,
		from 1 (almost never) to 7			physical fatigue accounted for unique variance in
		(always)			the burnout \( \mathbb{R} = \mathbb{R} = \mathbb{R} = \mathbb{R} \) burnout \( (R^2 = \mathbb{R} = \mathbb{R} = 24) \)
		Time points NR			0 •
Smith, 2017	CS	Sleep deprivation: self-	NR	Mental and physical illness:	Some physicians reported developing mental
J		reported via open-ended		self-reported via open-	illness (e.g., bipolated disorder, alcohol misuse) due
RoB: unclear		comments		ended comments	to tiredness and seess at work; others developed
					physical health prablems due to sleep deprivation
		Time points NR		Time points NR	poor eating habit and lack of exercise.

			ВМЈ Оре	en	3/bmjopen-2018 cted by copyrig
Study Risk of Bias (RoB)	Study design	Exposures or interventions Assessment measure and time points	Baseline	Outcomes  Assessment measure and time points	Associations between exposure and outcome  - Social Section
Starmer, 2016 RoB: low	CS	Sleep deprivation: <7 hours sleep in a typical 24-h period (self-reported)	27.7% sleep deprived	Burnout, satisfaction with career and life, balanced personal and professional commitments: Each on a 5-point Likert scale (strongly	≥7-h vs. <7-hd lee 1) Burnout ( vs. strong ly agree/agree): 26.4% vs. 39.6%, p<0.0 (vs. friedrich strongly agree/agree) (vs. strongly agree/agree) (vs. strongly agree/agree) (vs. strongly strongly strongly strongly strongly strongly strongly strongly) (vs. strongly stro
				agree to strongly disagree) Time points NR	p<0.05; balanced personal and professional commitments (2000) be professional and professional commitments (2000) commitments (
Tokuda, 2009 RoB: low	CS	Sleep hours/day: self- reported (continuous)  Time points NR (included weekday and weekends)	Mean±SD (range) sleep hours/day: 6±0.9 (3-8)	Burnout: MBI (Japanese) with a 7-point Likert scale: 0 (none) to 6 (every day); Job satisfaction: JHPSS with a 5-point Likert scale: 1 (strongly disagree) to 5 (strongly agree) Time points NR	Maximum likelihood estimates±SE:  1) Sleeping time to job satisfaction: group 0.990±0.458 = 0.331; ns for men; women 1.711±0.805 = 0.334;  2) Sleeping time to EEE: group -0.219 ±0.070, p=0.002; mei -0.215±0.082, p=0.009; ns for women.
Wada, 2010 RoB: unclear	CS	Sleep hours/day: Self- reported (continuous)  Assessed for past month when not completing overnight work	<5 hours: 8.7% men, 9.9% women; 5 to <6 hours: 32.3% men, 34.6% women; 6 to <7 hours: 46.0% men, 43.7% women; ≥7 hours: 13.0% men, 11.8% women.	Depression: QIDS-SR; Japanese score <5 (no symptoms) to >20 (very severe symptoms)  Assessed for past 7 days	1) Sleep house for hose with vs. without depressive suppressives: <5: 18.7% vs. 7.7% men, 20.5% vs. 8.7% we hen; 5 to <6: 33.7% vs. 32.2% men, 38.6% vs. 34% women; 6 to <7: 35.1% vs. 46.9% men; 31.8% vs. 45.1% women; 2) Association bet heen <5h sleep (vs. 6-7h) and depressive symptoms (OR (95% CI)): univariate 2.79 (1.96-3.95) for men, 2.65 (1.47-4.78) for women; multivariate (including age and workload)
		For peer revi	ew only - http://bmjopen.bm	.j.com/site/about/guideline	graphique de l

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	ncl
		time points		time points	udi

factors) 2.70 1.82 4.03) for men, 2.38 (1.11-5.10)

at Agence Bibliographique

for women. ថ្មី

cted by copyrig 5/bmjopen-2018-

alncludes studies of anesthetists, where these were physicians.

blincludes primary care physicians, internal medicine physicians, and general practitioners.

AM: morning; aMT6-s: melatonin metabolite; BA: before-after; CI: confidence interval; CBI: Copenhagen Burnout Inventory; CS: cross-sect Diagnostic and Statistical Manual of Mental Disorders; EE: emotional exhaustion; ER: emergency; ESS: Epworth Sleepiness Scale; GHQ: Geriging Realth Questionnaire; h: hour(s); ICU: intensive care unit; IL-8: interleukin-8; JDS: Job Diagnostic Survey; JHPSS: Japanese Hospital Physicians Satisfaction Scale; LAS: linear aparticles and logical control of the contr assessment scales; MBI: Maslach Burnout Inventory; MOSQ: Modified Occupational Stress Questionnaire; min: minute(s); NA: not applicate; not reported; ns: not statistically significant; OR: odds ratio; PA: personal achievement; PBM: Pines Burnout Measure; PE: professional efficacy; PM: afternoon; 👫 Frofile of Mood States; PSQI: Pittsburgh Sleep Quality Index; QIDS-SR: Quick Inventory Depressive Scale – Self-Reported; QOL: Quality of Life; RCT: randomized controlled in legal properties of the propert Pittsburgh Sleep Quality Index; QIDS-SR: Quick Inventory Depressive Scale – Self-Reported; QOL: Quality of Life; RCT: randomized controlled that RDAS: Revised Dyadic

		5/bmjopen-2018-02				
Performance and	safety c	outcomes related to fatigu	e or sleep loss among phy	rsicians in independent pra	ht, in 219	
Study	Study	gn Assessment measure and Baseline		Outcomes	Associatio between exposure and outcome	
Risk of Bias (RoB)	design			Assessment measure and	uding 2	
		time points		time points	<u> </u>	
Surgeons					or Se	
Uchal, 2005	RCT	Sleep deprivation from a 24-	Median (range) sleep hours:	Surgical performance:	Post call 💆 🏧 work:	
		h call shift vs. 8-h work;	1.5 (0-3) post-call vs. 6.5 (5-	laparoscopic surgical	1) Produc 📆 🚉 y: no difference in accuracy	
RoB: unclear		Sleep hours: self-reported	9) post-work, p<0.05;	simulator(Minimally Invasivs	error, tiss 🅰 🏝 age, leak rate;	
		(continuous);	Median ESS score: 7.0 post-	Surgical Trainer-Virtual	2) Proced receiveness: no difference in	
		Sleepiness: ESS (moderate:	call vs. 5.5 post-work, ns.	Reality) for product quality,	goal-directed	
		10-15, severe: ≥16)	•	procedure effectiveness	actions, o e per by g time.	
				·	xt & wn	
		Assessed post-call and post-		Assessed post-call and post-	llos erie	
		work		work	ide da	
Chu, 2011	СО	Sleep deprivation: self-	Of 4,047 procedures, 83	Surgical performance: CABG,	For 0-3 vs 5 .>6 hours of sleep: no	
•		reported hours, moderate	(2.1%) performed by	ACC	difference BBG or ACC.	
RoB: low		(3-6h) or severe (<3h)	severely sleep-deprived and		5 <del>5</del>	
		(5 21.)	1,595 (39.4%) moderately	Assessed during surgery	http://bmjo ng, Al trair	
		Assessed the night before	sleep-deprived surgeons		N t	
		surgery	and are managed and		p://bmjo Al train	
Ellman, 2004	СО	Sleep deprivation:	Of 6,751 procedures, 339	Surgical performance: CABG,	Sleep deprived vs. non-sleep deprived: no	
		performed a case starting	(5%) performed by sleep-	ACC	differences of CEBG or ACC.	
RoB: low		22:00 to 05:00, or ending	deprived surgeons	7.00	₫ 🚑	
NOD. 10W		22:00 to 07:30 and another	deprived surgeons	Assessed during surgery	sir <mark>C</mark>	
		case in the next 24-h		7.55c55cd ddinig sargery	com/ c	
Govindarajan,	СО	Sleep deprivation: treated	NR	Surgical performance:	Sleep dep <b>ត</b> ved vs. non-sleep deprived: no	
2015	-	patients from midnight to		duration of surgery	difference in diration of surgery, even after	
2013		07:00 and performed a		adiation of sargery	stratification by type of procedure.	
RoB: low		subsequent case on the				
NOB. IOW		same day			2025 ogies.	
Amirian, 2014	BA	17-h night call shift;	Naps pre-call: 11 (37%)	Surgical performance:	Pre- vs. post-car:	
, annian, 2014	υA	Sleep hours during the shift:	napped for median (IQR) 90	LapSimGyn laparoscopic	1) LapSimGyn 🚾 o difference in total time,	
RoB: high		Wrist-mounted Micro-Mini-	(58-128) min;	simulation for time, blood	blood loss, instrument path length, instrument	
NOD. HIGH		Motionlogger;	Median (IQR) sleep: 91 (62-	loss, instrument path;	angular path; napping did not affect	
		Sleepiness: KSS	123) min on the pre-call	D2 test of attention and	performance;	
		Siechiliess. VOS	night vs. 430 (329-449) on	concentration	_ · <del></del>	
			iligiit vs. 430 (323-443) 011	CONCENTRATION	<u>одг</u> ар	

			5/bmjopen-2018-				
					<del>5</del>		
Study	Study	Exposures or intervention		Outcomes	Associations between exposure and outcome		
Risk of Bias (RoB)	design	Assessment measure and time points	Baseline	Assessment measure and time points	967 o		
		Assessed on pre-call and on-	the on-call night, p<0.001;		2) D2 testamprovement in concentration,		
		call day; sleepiness assessed	Sleep on-call: 12 (40%) slept	Assessed on pre-call and on-	p<0.05. No changes in any other parameters;		
		during shift	for median (IQR) 98 (39-	call day	3) ns diffe ( in laparoscopic simulation tim		
			135) min;		in those who is but during the shift vs. not.		
			Significant development of		reig reig		
			sleepiness during shift		r 20 ner ate		
			(p<0.001), plateau score of		2018. ated to		
			7 at 04:00 to 08:00.		o te		
Gerdes, 2008	ВА	On-call shift;	Fatigue differential from	Psychomotor performance:	1) Pre- to क्रिक्टा ≰all: decrease in all measures o		
		Fatigue: questionnaire	pre- to post-call (range): 1-7	virtual ring transfer task for	psychomo of policiency (p<0.05, data NR)		
RoB: high		designed by Behrenz &	(units unclear);	gesture-level proficiency,	except el क्रुड्टिंग ime; no change in number of		
		Monga, 1999;	Sleep during call (range): 1-	hand movement	psychomological psychomologica		
		Sleep hours: self-reported	5h	smoothness, tool movement	(p<0.05, cat <b>2, 12, 12, 12, 12, 12, 12, 12, 12, 12, 1</b>		
		(continuous)		smoothness, elapsed time	2) Cognitive erpers increased exponentially as		
					fatigue ratings creased (R <sup>2</sup> =0.9219) and as		
		Assessed in 3 sessions pre-		Assessed in 3 sessions pre-	hours of steep beclined (R2=0.933).		
		and post-call		and post-call	ain jo		
Shanafelt, 2010	CS	Degree of fatigue as a	NR	Perceived recent major	1) Prevale ce de perceived recent major		
		contributor to errors (self-		medical errors (self-	medical eဋ္ဌ္လ္အဝr: <mark>န</mark> ွှိ 9%;		
RoB: unclear		reported)		reported)	2) Of those reparting an error, 6.9% listed		
					degree of atigge as the greatest contributing		
		Assessed for the past 3		Assessed for the past 3	factor. nii n/ o		
		months		months	te n		
Anesthesiologists <sup>a</sup>					) Chn		
Lederer, 2006	BA	24-h shift, on-call duty;	Mean±SD sleep: 4.1±1.7h;	Psychomotor performance:	Pre- vs. post-dမြို့y, mean±SD:		
		Sleep hours and	Number of interruptions:	reaction time, critical flicker	1) Psychometric testing: recognition reaction		
RoB: high		interruptions: self-reported;	0.8±1.1;	fusion, response measure,	time (ms) 439 ±50.8 vs. 480.3±58.9; motor		
		Tiredness: VAS from 0 (low)	Tiredness pre- vs. post-duty:	peripheral awareness;	reaction time (4s): 252.8±39.3 vs. 465.4±65.0		
		to 100 (high)	30.9±27.5 vs. 59.5±18.9,	Concentration ability: scale	total reaction the (ms): 690.8±73.4 vs.		
			p=0.01.	of 0 (low tiredness) to 100	746.5±113.7; cජූtical flicker fusion (Hz):		
		Assessed pre- and post-duty		(maximum tiredness)	29.0±2.3 vs. 28 23.7; response measure		
					(pixels): 647.8 26.7 vs. 598.3±138.1,		
				Assessed pre- and post-duty	ographic		

Study	Study	Exposures or intervention		Outcomes	Associations between exposure and outcom
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	967 ncl
		time points		time points	udi or
					periphera awageness task recognition time:
					58.9±59.20 s. 50.6±47.5;
					2) Concer <b>சூஷால்</b> ability: 26.4±23.5 vs. 56.3±23. <b>%ந்</b> . 200707.
Chang, 2013	CS	15-h in-house overnight call;	Median (IQR) ESS: 9 (9),	Psychomotor performance:	1) Aftern വുള്ള തുടലിലെ vs. pre-call: no
		Sleepiness pre-call: ESS ≥9;	64% scored ≥9;	reaction time; CCPT II; N-	difference described time, CCPT, N-back, of
RoB: unclear		Sleep hours: self-reported	Median (IQR) hours slept	back; HVLT (3 trials of 12	HVLT; to so so
		(continuous)	during shift: 1 (0-3).	words)	Morning 👸 🥳 e vs. post-call:
					1) No chaकिह्हिi≰auditory or visual reaction
		Sleepiness assessed pre-call,		Assessed at baseline and pre-	time; ne o
		sleep hours during call		and post-call	2) CCPT (tss.): No change in detectability response 歌魚山it reaction time,
					omissions <b>⊉com</b> Phissions;
					うじょう 3) N-back 多 accuracy: no change for auditory
					visual, or rearing visual, or rearing visual, or rearing visual, or rearing visual visual visual visual visual
					4) HVLT (tعد corغ : mean for trials 1-3: 48.6±7.
					vs. 41.5±9 (p .04); delayed recall: ns;
					5) No cormalation between ESS scores pre-cal
					or sleep divingshift and any measure of psychomogor performance.
Gander, 2000	CS	Nights of work-related sleep	NR	Risk of fatigue-related errors:	1) Risk of tige-related errors increased with
		disturbance: self-reported		questionnaire modelled after	increasing ights of work-related sleep
RoB: low		(continuous)		Gravenstein et al., 1990	disturbange: RP. 1.25, 95% CI: 1.06-1.49.
					June
		Assessed for the past 6		Assessed for the past 6	lune 9, 2
Condat 2017	CC	months	ND	months	<del> </del>
Saadat, 2017	CS	Sleep deprivation due to an overnight call shift	NR	Reaction time: PVT	Mean (SD) rea region time was slower post-cal (297.76 (83.75) vs. on a regular day (266.58
RoB: low		overinging can sillit		Assessed after an overnight	(38.35)), p=0.037.
-				call shift and the morning of	
				a regular (non-call) day	ence Bibliographique de l
					<u></u>

cted by copyrigh 6/bmjopen-2018-0

Study	Study	Exposures or intervention	Outcomes		Associations between exposure and outcome		
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	- ncl		
		time points		time points	ludi		
Gander, 2008	NC	Sleep loss across	≥2 hours sleep <baseline:< td=""><td>Psychomotor performance:</td><td>1) In fixed model analysis for reaction time</td></baseline:<>	Psychomotor performance:	1) In fixed model analysis for reaction time		
		consecutive working days or	8% of 24-h periods that	PVT	including <b>g</b> eep <b>g</b> time since waking, work hours		
RoB: unclear		on-call work: Wrist-	included day work vs. 14%		acute sleಈ 🎧 🚉 associated with slower mediar		
		mounted Actiwatch (Mini	that included day + call;	Assessed within 2 hours pre-	reaction $\mathbb{E}_{1,184} = 5.70$ , p<0.05; longer time		
		Mitter, Bend, Oregon, US),	Sleep hours: mean 0.6h less	and post-call	since wak 🖺 👸 ociated with poorer		
		sleep and duty diary	sleep when working day		performa the slowest 10%, F <sub>(1,185)</sub> =5.13,		
			shifts (p=0.014) and 0.8h		p<0.05; <b>c p o o o o o o o o o o</b>		
		Assessed over a 2-week	less sleep when working day		2) Reaction (In the second secutive work		
		period including a weekend	shifts + call (p=0.013) vs. off.		days: no c∯afe€ in pre-duty reaction times but		
		of rostered shifts or on-call			post-duty		
					median -3 4 6 0.001; decline in performance		
					across 10 <b>ឆ្នាំ ទ្រ</b> ប់ខ្មែន became progressively		
					steeper ball happer and post-duty, p=0.020.		
ER or ICU physician	ıs				<del>S</del> 5		
Sanches, 2015	CS	Acute sleep deprivation	Non-sleep deprived vs.	Psychomotor performance	Sleep deprived group vs. non-sleep deprived,		
		(<5h of night sleep after a	sleep deprived:	via Battery Test Reaction 5	mean±SD:		
RoB: high		night shift of 12h)		(v1): StimulTest, InstrucTest,	1) Instrucest: Forrect answers: 169.4 (16.0) vs.		
		Sleep hours: 7-day	PSQI >5: 0% vs. 33%, ns;	MovemTest; TP test of visual	148.3 (28万), p 0.070; wrong answers: ns;		
		Actigraphy via SenseWear®	ESS≥10: 11% vs. 67%	attention	perfection index (%): 99.6 (0.3) vs. 98.9 (1.3),		
		Pro2 Armband;	Sleep time (mean±SD) in		p=0.021; separase latency (sec/click): ns;		
		Sleepiness: ESS;	week before tests: duration	Assessed on morning after	2) StimulT≝st: orrect answers: 170.7 (21.9) vs.		
		Sleep quality: PSQI	and number of naps higher	night shift 8	145.1 (17年), p 0.022; wrong answers: ns;		
			in sleep deprived group, but		perfection index (%): ns; response latency		
		Assessed the week and	diurnal sleep hours lower,		(sec/click) 1.0€(0.1) vs. 1.24 (0.1), p=0.022;		
		night before the	428.6±30.1 vs. 375.8±55.9,		3) Moven est his for any parameter;		
		psychomotor tests	p=0.038;		4) TP: ometed symbols: 34.2±18.4 vs.		
			Sleep quality (mean±SD):		62.7 $\pm$ 44.0 $\mathbf{\ddot{g}}$ p=0 $\mathbf{\ddot{g}}$ 34; concentration index (%):		
			week before tests: 3.3±0.7		14.1±8.9 vs. 30 <u>x</u> 0±25.9, p=0.019; quality index		
			vs. 2.6±0.3, p=0.013;		(%): 13.8±8.6  29.2±26.4, p=0.031;		
			night before tests: 3.1±0.8		correct/wrong ymbols: ns;		
			vs. 1.9±1.0, p=0.020.		Correlations between sleep and tests:		
					1) TP for sleep bours nights 1-6: omitted		
					symbols: r=-0.66, p=0.011 for non-sleep-		
					graphique de		

			BMJ Oper	1	5/bmjopen-2018 cted by copyrig
Study Risk of Bias (RoB)	Study design	Exposures or intervention Assessment measure and	Baseline	Outcomes Assessment measure and	Obyright -2018 Associations between exposure and outcom
	_	time points		time points	u 7 or
					deprived as for sleep-deprived; concentration
					index (%) (\$\frac{1}{2}\$=-0.037 for sleep-
					deprived, \( \hat{H} \) s from non-sleep deprived; r=-0.359
					p=0.037 f <b>೫ ಸ್ಪ್ರಾ</b> group; no other significant
					correlation of the correlation o
					2) No cor and any the psychological tests.
• · · · ·					
Generalists <sup>b</sup>		0.01	4 4 6		te Xu
Harbeck, 2015	CS	24-hours on-call shift with	1) Sleep hours on a normal	Neurocognitive parameters:	Intrinsic a fertal size of the second attention and
		sleep disturbance: self-	day vs. following a 24-h	computerized attentional	vigilance Real Smilar on both occasions;
		reported number of sleep disturbances and hours of	shift: <2 hours: 0 vs. 5.9%;	test (vigilance, alertness); D2	Phasic ale
			2-4 hours: 5.9% vs. 47.1%;	letter cancellation test	shift: mean (2015), p
		sleep per night	4-6 hours: 11.8% vs. 35.3%; >6 hours: 82.4% vs. 11.8%	(divided attention); Trail	2.90 →
		Assessed before a normal	2) Number of sleep	Making Test (visual attention, task switching);	θ. · <del>(β</del>
		day shift, and after a 24-h	disturbances a normal day	Digit Span, Digit Symbol	Al 1
		on call shift	vs. following a 24-h shift:	Substitution Test, Weschler	rai mj
		on can sinit	0: 82.4% vs. 11.8%; 1: 11.8%	Memory Scale (memory	nin <mark>o</mark> g
			vs. 35.3%; 2: 5.9% vs. 47.1%;	functions)	9, n.b
			3: 0% vs. 5.9%; 4: 0% vs. 0%;		http://bmjopen.bmj.com/ on ) . ng, Al training, and similar t
			>4: 0% vs. 0%	Assessed before a normal	sin <mark>C</mark>
				day shift, and after a 24-h on	n/ o
				call shift	://bmjopen.bmj.com/ on Jur Al training, and similar tech
Mixed specialties o	r undefine	ed populations			c hr
Chen, 2008	CS	Sleepiness: ESS score ≥11	Mean±SD ESS score:	Impact on work and personal	1) Impact core correlated with ESS, r=0.31,
			7.8±4.0, range: 0-20, 23%	life: Impact Questionnaire	p<0.05; <b>9</b> 20
RoB: high		Time points NR	had scores ≥11.	with a 5-point Likert scale	2) ESS score was higher among physicians wh
				from 1 (strongly agree) to 5	agree/strongly gree vs. other response:
				(strongly disagree)	written an incerrect order: 8.8 vs. 7.3, p=0.02
					might fall asle while examining a patient:
				Time points NR	13.2 vs. 7.7, p=0.001; look forward to sleepin
					at grand round 10.4 vs. 7.4, p=0.002;
		For neer revie	w only - http://bmjopen.bmj.	com/site/about/quidelines x	ographique de l

cted by copyrigh 6/bmjopen-2018-(

Assessment measure and time points  Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)	Mean±SD (range) score: 2.30±1.00 (1-6)	Assessment measure and time points  Work ability: Work Ability Index on scale from 1 (could not work at all) to 10 (best	3) No difference in ESS score for those who agree/strongly or gree vs. other response: work is unaffected by sleep loss and fatigue, thinking is unaffected by sleep loss, sleep loss and fatigue after making medical decisions, have heard of which was being making medical errors due to sleep loss and stigue, never make errors in prescription post-call days, have made medical effects because of sleep loss and fatigue; and work secause of sleep loss and fatigue
Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)	Mean±SD (range) score:	Work ability: Work Ability Index on scale from 1 (could	3) No difference in ESS score for those who agree/strongly agree vs. other response: work is unaffected by sleep loss and fatigue, thinking is unaffected by sleep loss, sleep loss and fatigue after ymaking medical errors due to sleep loss and betigue, never make errors in prescription post-call days, have made medical effects secause of sleep loss and fatigue;  4) Higher as a prescription of the prescr
Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)	Mean±SD (range) score:	Index on scale from 1 (could	agree/strongly agree vs. other response: work is unaffect of the sleep loss and fatigue, thinking is unaffect of the sleep loss, sleep loss and fatigue after the making medical errors due to sleep loss and being making medical errors due to sleep loss and being post-call days, have made medical errors in prescription post-call days, have made medical errors in prescription of post-call days, have made medical errors in prescription of post-call days, have made medical errors in prescription of post-call days, have made medical errors in prescription of post-call days, have made medical errors in prescription of post-call days, have made medical errors and fatigue;  4) Higher as a prescription including personal and work approach and indirect effect on work ability (R <sup>2</sup> 0.1 195% CI: -0.122, -0.031,
Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)	Mean±SD (range) score:	Index on scale from 1 (could	is unaffected by sleep loss and fatigue, thinking is unaffected by sleep loss, sleep loss and fatigue after by medical decisions, have heard of the by making medical errors due to sleep loss and betigue, never make errors in prescription by post-call days, have made medical effects because of sleep loss and fatigue;  4) Higher by secause of sleep loss and fatigue;  and work regression including personal and work regression including personal and work regression indirect effect on work ability (R <sup>2</sup> -0.1 <sup>1</sup> / <sub>2</sub> 95% CI: -0.122, -0.031,
Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)	Mean±SD (range) score:	Index on scale from 1 (could	is unaffected by selep loss, sleep loss and fatigue af the property medical decisions, have heard of control by medical decisions, have heard of control by making medical errors due to sleep loss and being post-call days, have made medical errors because of sleep loss and fatigue;  4) Higher control by medical by impact score in multivation because predicted by impact score in multivation because of sleep loss and fatigue;  1) On-call array and an indirect effect on work ability (R <sup>2</sup> -0.11, 95% CI: -0.122, -0.031,
Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)	Mean±SD (range) score:	Index on scale from 1 (could	fatigue af recommended programs of the process of
Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)	Mean±SD (range) score:	Index on scale from 1 (could	heard of heart sharing medical errors due to sleep loss and be being post-call days, have made medical errors because of sleep loss and fatigue;  4) Higher and heart sharing personal and work representations of factors: β=0.11, p=0.005.  1) On-call and yellow an indirect effect on work ability (R <sup>2</sup> -0.11, 95% CI: -0.122, -0.031,
Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)	Mean±SD (range) score:	Index on scale from 1 (could	sleep loss and Batigue, never make errors in prescription for post-call days, have made medical effects secause of sleep loss and fatigue; and fati
Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)	Mean±SD (range) score:	Index on scale from 1 (could	prescription post-call days, have made medical effects secause of sleep loss and fatigue; and response predicted by impact score in multivation feet effects and work regression including personal regression including personal and work regression including personal and work regression including personal regression
Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)	Mean±SD (range) score:	Index on scale from 1 (could	medical experience of sleep loss and fatigue;  4) Higher as a period of the second in multival and regression including personal and work are first and an indirect effect on work ability (R <sup>2</sup> -0.1½,95% CI: -0.122, -0.031,
Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)	Mean±SD (range) score:	Index on scale from 1 (could	fatigue; and a display of the second of the
Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)	Mean±SD (range) score:	Index on scale from 1 (could	4) Higher (See Secretary Secretary) And Higher (See Secretary Secretary) And Higher (Secretary Secretary) And Higher (Secretary Secretary Secretary) And Higher (Secretary Secretary Secr
Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)	Mean±SD (range) score:	Index on scale from 1 (could	in multivariant features in multivariant features in multivariant features from the features in multivariant features from the feature of th
Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)	Mean±SD (range) score:	Index on scale from 1 (could	and work and an indirect effect on work ability (R <sup>2</sup> 0.1 $\frac{1}{10}$ 95% CI: -0.122, -0.031,
Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)		Index on scale from 1 (could	1) On-call 30 y and an indirect effect on work ability (R 0.15,95% CI: -0.122, -0.031,
Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)		Index on scale from 1 (could	ability (R $\frac{\mathbf{Q}}{5}$ 0.1 $5$ 95% CI: -0.122, -0.031,
scale from 1 (never) to 6 (every night)	2.30±1.00 (1-6)	·	<b>D</b> 🛬
(every night)		not work at all) to 10 (best	p<0.001) farough two mediators (work
		work ability)	interferer
			2) Sleeping prolims inversely associated with
Assessed in 2006		Assessed in 2010	work ability, $\beta = 0.29$ , p<0.001.
Sleep hours: self-reported	Insufficient rest: 32.5%;	Self-reported medical	1) Prevaleace at medical incidents (% (95%
(continuous)	Daytime sleepiness: 3.5%;	incidents: 4-point scale from	CI)): sleep (epigyed (26.8% (24.2, 29.4)) vs. not
Sleepiness and sleep	Insomnia: 20.0%;	1 (never) to 4 (often)	(15.2% (1ﷺ7, 15.7)), p<0.01; insomnia (24.8%
difficulties: 5-point scale	Sleep time (mean±SD min):		(21.6, 28.2) insomnia vs. not (17.6% (16.2,
	279.8±60.9	Assessed for the past month	19.0)), p< 201; 56h sleep (18.3% (16.8, 19.8))
Insomnia: ≥3 sleep			vs. <6h (2 <b>2</b> 7% <b>9</b> 18.8, 24.6)), p=0.03;
difficulties			2) Predicters of medical incidents in
			multivaria medel including personal and
Assessed for the past month			work-related factors (OR (95% CI)): lacking res
			due to sleep derivation vs. not (1.65 (1.33-
			2.04)), p<0.01) <b>g</b> insomnia vs. not (1.45 (1.16-
			1.82), p<0.01); os for sleep hours.
	Sleepiness and sleep difficulties: 5-point scale from 1 (never) to 5 (always); Insomnia: ≥3 sleep difficulties	Sleepiness and sleep Insomnia: 20.0%; difficulties: 5-point scale Sleep time (mean±SD min): from 1 (never) to 5 (always); Insomnia: ≥3 sleep difficulties	Sleepiness and sleep Insomnia: 20.0%; 1 (never) to 4 (often) difficulties: 5-point scale Sleep time (mean±SD min): from 1 (never) to 5 (always); 279.8±60.9 Assessed for the past month Insomnia: ≥3 sleep difficulties

			BMJ Oper	1	5/bmjopen-2018-
Study	Study	Exposures or intervention		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	
	_	time points		time points	967 o
Sexton, 2001	CS	Fatigue as a factor	NR	Performance effectiveness	1) "When that ig wed, I perform effectively during
		impacting performance		measured by 1 question:	critical pheses of operations/patient care":
RoB: high				agree, neutral, disagree	Anestheti 🖁 🚜 agree; 15% neutral; 38%
-		Time points NR			disagree; 🕉 🙀 🖺
				Time points NR	Surgical: 전描 agree; 12% neutral; 18%
					disagree. e 20
Shirom, 2006	CS	Tiredness and exhaustion:	NR	Quality of care: Adapted 15-	1) Quality 12 ame positively predicted by
		SMBM Physician Fatigue		item SERVQUAL with a 5-	fatigue in a model incorporating several other
RoB: low		Subscale on a 7-point scale		point Likert scale from 1	compone 📆 🕏 🕏 purnout, β=0.17, p<0.05.
		from 1 (almost never) to 7		(very small extent) to 5 (very	and and
		(always)		large extent)	I dat
		Time points NR		Time points NR	a mir
Smith, 2017	CS	Sleep deprivation: self-	NR	Perceived competence: self-	Some physicians indicated that continual
		reported via open-ended		reported via open-ended	tiredness and shaustion led to concerns that
RoB: moderate		comments		comments	it would affect their competence; some felt
					that profesion performance was
		Time points NR		Time points NR	compromised at times of physical and mental
					fatigue. 🚡 💆
Tanti, 2017	CS	Fatique: questionnaire on	NR	Prescribing errors:	Perceptio of the contribution of fatigue to
		contributors to prescribing		questionnaire on	prescribing errers differed by physician type
RoB: high		errors, with a 5-point Likert		contributors to prescribing	(p<0.05): 34% of community doctors, 96%
		scale (very high to very low		errors, with a 5-point Likert	hospital dectors, 8% of office-working doctors
		association)		scale (very high to very low	perceived vest high or high association
				association)	between atigue and prescribing errors.
		Time points NR			0 ,
				Time points NR	2025 gies.

<sup>&</sup>lt;sup>a</sup>Includes studies of anesthetists, where these were physicians.

ACC: aortic cross-clamp time; BA: before-after; CABG: cardiopulmonary bypass time; CCPT II: Connor's Continuous Performance Test II; CI: configence interval; CO: cohort; CS: cross-sectional; ER: emergency; ESS: Epworth Sleepiness Scale; h: hour(s); HVLT: Hopkin's Verbal Learning Task; Hz: Hertz; ICU: intensive care until IQR: interquartile range; KSS: Karolinska Sleep Scale; min: minutes; ms: millisecond(s); N-back: Dual N-back test; NA: not applicable; NR: not reported; ns: not statistically significant; OR: odds ratio; PSQI: Pittsburgh Sleep Quality Index; PVT: Psychomotor vigilance Performance Task; RR: risk ratio; RCT: randomized controlled trial; RoB: Risk of Bias D: standard deviation; SE: standard error; SERVQUAL: Service Quality Measure; SMBM: Shirom-Melamed Burnout Measure; TP: Toulouse-Piéron test; TS: time series; US: nited States of America; vs.: versus

<sup>&</sup>lt;sup>b</sup>Includes primary care physicians, internal medicine physicians, and general practitioners.

cted by copyright, 5/bmjopen-2018-02

 Patient outcomes related to fatigue or sleep restriction among physicians in independent practice

Study	Study	Exposures		Outcome Measures	Associaten set tween exposure and outcom
Risk of Bias (RoB)	design	Intervention or assessment	Baseline	Assessment scale and time	- id on
		scale and time points		points	on 21
Surgeons					or Se
Chu, 2011	СО	Sleep deprivation: moderate	Of 4,047 procedures, 83	Chart review: mortality, surgical	1) 0-3 vs இ ஞ 😸 . >6 hours of sleep: No
		(3-6 h) or severe (<3-h) sleep	(2.1%) performed by	complications, length of stay	differencਵੇਂ ਲਿੰ i ਵਿcidence of mortality, incidenc
RoB: low		deprivation the night before	severely sleep-deprived,		of 10 ma இத்திறிications (except septicemia
		surgery (self-reported hours)	1,595 (39.4%) by	Assessed during and post-	3.6% vs. <b>ૄ ૾</b> ઽઽ. 0.8%, p=0.03), ICU length o
			moderately sleep-	surgery	stay; in-hos im length of stay (days): 7.0 vs.
			deprived surgeons		6.0 vs. 7. <b>5 p. 5</b> 001.
Ellman, 2004	СО	Sleep deprivation: performed	Of 6,751 procedures,	Chart review: mortality, surgical	1) Sleep ( see deprived: no
		a case starting 22:00 to 05:00,	339 (5%) were	complications, length of stay	difference நெற்றாகிortality, need for blood
RoB: low		or ending 22:00 to 07:30 and	performed by sleep		productsနိုင်င်ကြွှဲlications (operative,
		performed a subsequent case	deprived surgeons	Assessed during and post-	neurologic, 🚉 😎 I, infectious, pulmonary), in-
		in the next 24-h		surgery	hospital 🖫 📆 🕏 of stay.
Govindarajan,	СО	Sleep deprivation: treated	NR	Chart review: mortality, surgical	1) Sleep prized vs. non-sleep deprived: no
2015		patients from midnight to		complications, readmission,	difference in Rortality, surgical complications
		07:00 and performed a		length of stay	readmiss ns ithin 30 days, or length of star
RoB: low		subsequent case on the same			ain jo
		day		Assessed during and post-	jopen. ining,
				surgery	a b
Rothschild, 2009	СО	Sleep deprivation: daytime	NR	Chart review: frequency of	1) Post-nighttime vs. control: no difference in
		procedures following an		adverse surgical complications	number $\overset{lack}{\underline{\underline{\mathbf{g}}}}$ pr $\overset{lack}{\underline{\underline{\mathbf{g}}}}$ edures with complications,
RoB: low		overnight procedure;			total number of complications, preventable
		Sleep opportunity: 0-6h, <6h		Assessed during and post-	complications type of complications;
				surgery	2) Operating room procedures with
					complica OR (95% CI): 8.5% for 0-6h
					sleep vs. <b>4.</b> 1% <b>2</b> or >6h sleep, 2.70 (1.13-6.48),
					p=0.03; <b>% S</b>
					3) All procedures with complications, OR (959
					CI): 6.2% for 66h sleep vs. 3.4% for >6h sleep
					1.72 (1.02-2.8 <b>ặ</b> ), p=0.04.
					Ö
					Si Di
					<u>ö</u> 0
					Bibliographique de l
					hiq
					ue
		For peer review	only - http://bmjopen.br	nj.com/site/about/guidelines.xl	ntml 🔐

			ВМЈ Ор	en	√bmjo cted b
					5/bmjopen-2018-i cted by copyrigh
Study	Study	Exposures		Outcome Measures	Associations 腔tween exposure and outcom
Risk of Bias (RoB)	design	Intervention or assessment	Baseline	Assessment scale and time	967
		scale and time points		points	udi
Schieman, 2007	СО	Fatigue: surgeon billed for	Of 270 procedures, 22	Chart review: surgical	1) Fatigued vsknon-fatigued surgeons: no
		clinical work after 22:00 the	(8%) were performed by	complications, length of stay,	differenc <b>ें</b> in i <b>क्रा</b> ra- or post-operative
RoB: low		night before surgery	fatigued surgeons	mortality, cancer recurrence	complica ( in-hospital length of stay, in-hospital length of stay) in-hospital length of stay.
				Assessed during and post-	ber 20 eigner relate
				surgery	20 nen atec
Vinden, 2014	СО	Sleep deprivation (at risk):	Of 94,183 surgeries,	Chart review: conversion to	1) At risk surgeon: no difference
		surgeon worked 00:00 to	2,078 (2.2%) were	open procedure (from	in incider conversion to open procedure
RoB: low		07:00 and performed surgery	performed by surgeons	laparoscopic), iatrogenic	iatrogeni இத்துத் ies, mortality, in either
		07:00 to 18:00	who were 'at risk'	injuries, mortality	univariate analyses.
				Assessed during and post-	d fr
				surgery	mir SEC
Obstetricians					htt s) .
Rothschild, 2009	СО	Sleep deprivation: daytime	NR	Chart review: frequency of	1) Post-nighttime vs. control: no difference in
		procedures following an		adverse obstetric complications	number ब्र्ने prब्दुedures with complications,
RoB: low		overnight procedure;			total con මු lications, preventable
		Sleep opportunity: 0-6h, <6h		Assessed during and post-	complications type of complications;
				delivery	2) No association between sleep deprivation
					and propartion of procedures with
					complica ons on difference for 0-6h vs. >6h
					of sleep குற்றர் unity.
: confidence interv	al; CO: col	nort; h: hours; NR: not reported; (	DR: odds ratio; RoB: Risk of I	Bias; SD: standard deviation; US: Uni	complications of sleep apportunity. ited States of America; vs.: versus
		For peer review	only - http://bmjopen.br	nj.com/site/about/guidelines.xh	2025 at Agence Bibliographique de l gies.

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

#### **Dichotomous outcomes**

Outcome or subgroup	Number	Number of	Pooled risk	Heterogeneity	
	of studies	participants	ratio (95% CI)	Р	l <sup>2</sup>
1.1 Patient mortality	5	60,436	0.98 (0.84, 1.15)	0.73	0%
1.2 Intra-operative	3	19,798	suppressed	0.007	82%
complications					
1.2.1 Surgical procedure	3ª	14,896	suppressed	<0.001	88%
1.2.2 Obstetric procedure	<b>1</b> <sup>a</sup>	4,902	suppressed	NA	NA
1.3 Post-operative	5	60,201	0.99 (0.95, 1.03)	0.45	0%
complications					

<sup>&</sup>lt;sup>a</sup> Rothschild, 2009 is represented in both analyses

#### **Continuous outcomes**

Outcome or subgroup	Number of	Number of	Pooled mean	Heterog	eneity
	studies	participants	difference (95% CI)	Р	l <sup>2</sup>
1.4 Operating time (minutes)	4	50,046	-0.14 (-1.60, 1.33)	0.70	0%
1.5 Length of hospital stay (days)	4	50,046	suppressed	<0.001	86%
1.5.1 Cardiac surgeries	2	10,798	suppressed	0.01	84%
1.5.2 Elective surgeries	1	38,978	suppressed	NA	NA
1.5.3 Anterior resection for anal	1	270	suppressed	NA	NA
cancer		<b>V</b> ,			

CPBT: cardiopulmonary bypass time; NA: not applicable

### 1.1 Patient mortality

	Sleep de	prived	Not sleep de	eprived		Risk Ratio		Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI	Year	M-H, Random, 95% CI
Ellman 2004	17	339	296	6412	10.4%	1.09 [0.67, 1.75]	2004	<u>+</u> -
Schieman 2008	0	22	3	248	0.3%	1.55 [0.08, 29.03]	2008	<del></del>
Chu 2011	47	1678	80	2369	18.8%	0.83 [0.58, 1.18]	2011	<del></del>
Vinden 2014	2	2078	7	8312	1.0%	1.14 [0.24, 5.50]	2014	<del></del>
Govindarajan 2015	224	19489	222	19489	69.5%	1.01 [0.84, 1.21]	2015	•
Total (95% CI)		23606		36830	100.0%	0.98 [0.84, 1.15]		•
Total events	290		608					
Heterogeneity: Tau <sup>2</sup> =	: 0.00; Chi²	= 1.25, d	f = 4 (P = 0.8)	7); I <sup>2</sup> = 0%			F	1.01 0.1 10 100
Test for overall effect:	Z = 0.23 (F	P = 0.82)					U	0.01 0.1 1 10 100 Favors sleep deprived Favors non-sleep deprived

#### Sensitivity analysis using highest possible number of events for Vinden 2014

	Sleep de	prived	Not sleep d	eprived		Risk Ratio		Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI	Year	M-H, Random, 95% CI
Ellman 2004	17	339	296	6412	13.9%	1.09 [0.67, 1.75]	2004	<del>-</del>
Schieman 2008	0	22	3	248	0.4%	1.55 [0.08, 29.03]	2008	<del> </del>
Chu 2011	47	1678	80	2369	23.3%	0.83 [0.58, 1.18]	2011	<del></del>
Vinden 2014	5	2078	7	8312	2.6%	2.86 [0.91, 8.99]	2014	<del></del>
Govindarajan 2015	224	19489	222	19489	59.8%	1.01 [0.84, 1.21]	2015	<b>†</b>
Total (95% CI)		23606		36830	100.0%	1.00 [0.83, 1.21]		<b>•</b>
Total events	293		608					
Heterogeneity: Tau <sup>2</sup> =	0.01; Chi²	= 4.50, d	f = 4 (P = 0.34)	4); $I^2 = 119$	6			
Test for overall effect:	Z = 0.03 (F	P = 0.98)						0.01 0.1 1 10 100 Favors sleep deprived Favors non-sleep deprived

BMJ Open: first published as 10.1136/bmjopen-2018-021967 on 21 September 2018. Downloaded from http://bmjopen.bmj.com/ on June 9, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

## 1.3 Post-operative complications

	Sleep de	prived	Not sleep de	eprived	Risk Ratio			Risk Ratio		
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI		M-H, Random, 95% CI		
Chu 2011	259	1678	404	2369	7.8%	0.91 [0.78, 1.04]				
Ellman 2004	29	339	480	6412	1.2%	1.14 [0.80, 1.64]		<del></del>		
Govindarajan 2015	3527	19489	3543	19489	89.4%	1.00 [0.95, 1.04]				
Schieman 2008	12	22	164	248	1.0%	0.82 [0.56, 1.22]		<del></del>		
Vinden 2014	14	2031	72	8124	0.5%	0.78 [0.44, 1.38]				
Total (95% CI)		23559		36642	100.0%	0.99 [0.95, 1.03]		•		
Total events	3841		4663							
Heterogeneity: Tau <sup>2</sup> = 0.00; Chi <sup>2</sup> = 3.69, df = 4 (P = 0.45); $I^2$ = 0%										
Test for overall effect:				0.2	0.5 1 2 5 Favors sleep deprived Favors non-sleep deprived					

#### 1.4 Operating time (minutes)

	Sleep deprived			Not sleep deprived				Mean Difference		Mean Difference		
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	Year	IV, Random, 95% CI		
Ellman 2004	107.7	55.24	339	107.4	56.05	6412	5.9%	0.30 [-5.74, 6.34]	2004	+		
Schieman 2008	114	226.6	22	148	226.6	248	0.0%	-34.00 [-132.80, 64.80]	2008	· · · · · · · · · · · · · · · · · · ·		
Chu 2011	110.2	119.13	1678	114	120.88	2369	3.8%	-3.80 [-11.30, 3.70]	2011	<del> </del>		
Govindarajan 2015	156	75.56	19489	156	80	19489	90.2%	0.00 [-1.54, 1.54]	2015	· ·		
Total (95% CI)			21528			28518	100.0%	-0.14 [-1.60, 1.33]		•		
Heterogeneity: Tau² = Test for overall effect				P = 0.70	); I² = 0%					-100 -50 0 50 100 Favors sleep deprived Favors non-sleep deprived		



45 46 47

Page 85 of 87		BMJ Open BMJ open				
by copyrig						
Section/topic	#	Checklist item the control of the co	Reported on page #			
TITLE		ding 2				
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1			
ABSTRACT		use:				
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; state gibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; concessions and implications of key findings; systematic review registration number.	2			
INTRODUCTION		Doy				
Rationale	3	Describe the rationale for the review in the context of what is already known.	4-5			
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5			
METHODS		m BES;				
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if a like, provide registration information including registration number.	5			
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years encourage, publication status) used as criteria for eligibility, giving rationale.	6			
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors identify additional studies) in the search and date last searched.	5-6			
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that income a search strategy for at least one database, including any limits used, such that income a search strategy for at least one database, including any limits used, such that is could be repeated.	Supplementary file 1			
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and included in the meta-analysis).	6-7			
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.	7			
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assum ions and simplifications made.	7			
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.	7			
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	8			
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures for consistency (e.g., I <sup>2</sup> ) for each meta-analysis.	8			
4 5 6		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml				



45 46 47

		BMJ Open	Page 86 of 8
Appendix 1. PRIS	BMJ Open  BMJ Open  klist		
3 4 Section/topic 5	#	Checklist item	Reported on page #
6 Risk of bias across studies 7	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication )	8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression	8
1 RESULTS		relation to the second	
12 13 13	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reason exclusions at each stage, ideally with a flow diagram.	8, Figure 1
15 Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follows and provide the citations.	8-11, Table 1, Supplementary file 2
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item ).	11, Supplementary file 3
Results of individual studies  23  24	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.	p. 12-18; Supplementary file 4; figures 2-4
6 Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consecutive.	p. 12-18, figures 2-4
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Not applicable
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [55] [15] [15]	Supplementary file 5
DISCUSSION		olog	
34 Summary of evidence 35	24	Summarize the main findings including the strength of evidence for each main outcome; consider the relevance to key groups (e.g., healthcare providers, users, and policy makers).	18-19
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).	19-20
39 Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for Eture research.	20
FUNDING	<u> </u>	g	
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of unders for the systematic review.	21
<del>                                   </del>		<del></del>	

de l

\*\*September 2018. Downtoaded Home No. 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyss. Including for uses Related to use and data mining. At training, and similar technologies.

# **BMJ Open**

# The impact of fatigue and insufficient sleep on physician and patient outcomes: A systematic review

Journal:	BMJ Open				
Manuscript ID	bmjopen-2018-021967.R3				
Article Type:	Research				
Date Submitted by the Author:	07-Aug-2018				
Complete List of Authors:	Gates, Michelle; University of Alberta, Pediatrics Wingert, Aireen; University of Alberta, Pediatrics Featherstone, Robin; University of Alberta, Pediatrics Samuels, Charles; Centre for Sleep and Human Performance Simon, Christopher; Canadian Medical Association Dyson, Michele; University of Alberta, Pediatrics				
<b>Primary Subject Heading</b> :	Occupational and environmental medicine				
Secondary Subject Heading:	Evidence based practice, Health services research, Medical education and training, Occupational and environmental medicine				
Keywords:	Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, SLEEP MEDICINE				

SCHOLARONE™ Manuscripts

mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

The impact of fatigue and insufficient sleep on physician and patient outcomes: A systematic review

Michelle Gates, Aireen Wingert, Robin Featherstone, Charles Samuels, Christopher Simon, Michele P Dyson

**Michelle Gates**, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-486C, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Project coordinator

**Aireen Wingert**, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-490B, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Project coordinator

**Robin Featherstone**, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-486D, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Research librarian

**Charles Samuels**, Centre for Sleep and Human Performance, Suite 106, 51 Sunpark Drive SE, Calgary, Alberta, Canada, T2X 3V4
Physician and medical director

**Christopher Simon**, Canadian Medical Association, 1867 Alta Vista Drive, Ottawa, Ontario, Canada, K1G 5W8

Advisor, physician health

Michele P Dyson, Alberta Research Centre for Health Evidence (ARCHE), Edmonton Clinic Health Academy 4-474, University of Alberta, 11405-87 Avenue NW, Edmonton, Alberta, Canada T6G 1C9 Assistant professor

Correspondence to: MP Dyson mdyson@ualberta.ca

Word count (main text): 5,929

Objectives: For physicians in independent practice, we synthesized evidence on the (a) impacts of insufficient sleep and fatigue on health and performance, and patient safety; (b) effectiveness of interventions targeting insufficient sleep and fatigue.

Design: We systematically reviewed online literature. After piloting, one reviewer selected studies by title and abstract; full texts were then reviewed in duplicate. One reviewer extracted data; another verified a random 10% sample. Two reviewers assessed risk of bias. We pooled findings via metaanalysis when appropriate, or narratively.

Data sources: We searched Medline, Embase, PsycINFO, CINAHL and PubMed for published studies in April 2016; Medline was updated in November 2017. We searched Embase for conference proceedings, and hand-searched meeting abstracts, association and foundation websites.

Eligibility criteria for selecting studies: English or French language primary research studies published from 2000-2017 examining the effect of fatigue or sleep-related exposures or interventions on any outcome among physicians in independent practice and their patients.

Results: Of 16,154 records identified, we included 47 quantitative studies of variable quality. 28 studies showed associations between fatigue or insufficient sleep and physician health and well-being outcomes. 21 studies showed no association with surgical performance, and mixed findings for psychomotor performance, work performance, and medical errors. We pooled data from six cohort studies for patient outcomes. For sleep deprived versus non-sleep deprived surgeons, we found no difference in patient mortality (n = 60,436, RR 0.98, 95% CI 0.84 to 1.15,  $I^2 = 0\%$  (P = 0.73)) nor postoperative complications (n = 60,201, RR 0.99, 95% CI 0.95 to 1.03,  $I^2 = 0\%$  (P = 0.45)). The findings for intraoperative complications and length of stay were considerably heterogeneous.

**Conclusions:** Fatigue and insufficient sleep may be associated with negative physician health outcomes. Current evidence is inadequate to inform practice recommendations.

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text and

- The review was informed by the methods outlined by Cochrane and is reported according to the
   Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.
- The review was limited by the quality of the included studies, which was often poor. Confidence in our conclusions may be weakened due to multiple comparisons.
- We have focused on evidence from high income countries; our findings may not be generalizable to other settings.



The working hours of physicians have been a topic of debate for many years.[1] Beginning in the late 1980s, evidence indicating that medical resident fatigue could negatively impact their cognitive functioning and performance, resulting in an increased risk of medical error, began to accumulate.[2] In response, by the early 2000s physicians' regulatory bodies worldwide began to take action toward restricting the work hours of medical residents and ensuring adequate time for recovery between shifts.[3-5] Since their implementation in the United States by the Accreditation Council for Graduate Medical Education (ACGME), the impact of work hour regulations has been widely researched. Still, evidence for impacts on patient care, resident training and wellbeing remains equivocal.[6-9] This is likely because work hours are only one of many contributors to fatigue and physician wellbeing. In fact, the ACGME has recently reversed the 2011 changes that limited resident work hours to 16 hours per shift and the requirement for 8 hours of time off between shifts. This decision was made in favour of promoting "flexibility" for residency training program work hours and scheduling.

The focus on medical trainees has left physicians in independent practice as a relatively neglected group in research and policy. In Canada, there is no concrete regulation on the hours or patterns in which physicians choose to work.[10] In the absence of clear policies, physicians trained under traditional systems may find it difficult to work shorter hours or take more frequent breaks.[1] Indeed, more than 40% of practicing physicians in the United States work in excess of 80 hours per week.[11] While long work hours remain a cultural norm in medicine, in comparable high-risk industries (e.g., aviation), work patterns and work hours are tightly regulated.[12] The need for similar evidence-based policies in medicine has become a topic of increased interest. Exemplar of this, an evidence-based guideline for fatigue risk management in emergency medical services,[13] informed by a comprehensive set of systematic reviews, has recently been published. For physicians, it has been argued that there is a need to adapt healthcare systems and provide support in identifying the signs of fatigue and mitigating its risks.[1]

Besides potentially affecting patient outcomes, fatigue can impact the health and wellbeing of physicians themselves. Burnout, just one outcome related to fatigue, has been described as epidemic among physicians[14-16] and ultimately affects recruitment and retention of physicians both in community and acute care settings. While the effect of physician wellbeing on the sustainability of healthcare systems has recently received increased attention,[17] evidence-based solutions to burnout

remain relatively elusive.[18] What is clear, is that comprehensive organisational-level efforts are necessary to fully address the issue.[19] Research addressing the factors that influence burnout and overall physician wellness is needed to inform system- and individual-level strategies.[20, 21] To date, evidence of the effects of fatigue and the role of chronic insufficient sleep on physicians in independent practice has not been synthesized, making it unclear what gaps in knowledge remain unaddressed.

Given this void, we undertook a systematic review focusing broadly on primary research relevant to the Canadian context as a fundamental starting point to examine the effects of fatigue and chronic insufficient sleep on physicians in independent practice, and on interventions to combat these effects. Our review was guided by the following research questions: Among physicians in independent practice, (1) what are the impacts of fatigue and chronic insufficient sleep on physician health, physician performance, and patient safety; and (2) what is the effectiveness of interventions that target fatigue and chronic insufficient sleep, in terms of improving physician and patient outcomes?

#### **METHODS**

#### **Review conduct**

The conduct of this systematic review was guided by Cochrane standards.[22] The research team convened to plan the key research questions and methodology but did not register a formal protocol. The findings are reported in adherence with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) Statement.[23] Ethical approval was not required for this study.

#### **Patient involvement**

Patients were not involved.

## Literature search

An information specialist developed a search strategy that included concepts related to physicians, fatigue and sleep. On 13 April 2016 we searched the following online databases with coverage in the biomedical sciences and psychology: Medline, Embase, PsycINFO, CINAHL and PubMed, limited to English and French language articles published from 2000 to 2016. We updated the Medline search in November 2017, as this database offered the highest precision. Though fatigue among physicians is not a new phenomenon,[2] we limited our search to articles published post-2000 to include studies relevant to current physician practice. Work hour limitations have existed in European countries since 1993, but

#### Inclusion criteria

Primary studies (quantitative or qualitative) of fatigue- or sleep-related exposures or interventions among physicians in independent practice were eligible for inclusion. We included physicians practicing in any medical specialty and in any healthcare setting within a high income country,[25] to identify practices comparable to the Canadian setting. Studies including physicians-in-training were included only if data for physicians in independent practice could be isolated. Exposures of interest included fatigue, insufficient sleep, or sleepiness. We also included studies of any intervention that aimed to reduce fatigue or sleep loss with any comparator (or no comparator). All reported outcomes, measured at any time, were eligible for inclusion.

We excluded commentaries, letters, editorials and dissertations. Systematic reviews, health technology assessments, economic evaluations and practice guidelines were excluded, although the reference lists of these as well as the included studies were scanned for potential primary studies. Studies that focused solely on physicians-in-training (e.g., trainees, residents, fellows, interns, medical students, junior doctors, registrars) were ineligible. To maintain the focused scope of the review, we excluded work hours, work load, and any other exposure or intervention that was indirectly related to fatigue or sleep.

# Study selection

The study team piloted the selection criteria, which were then applied by two independent reviewers following a two-phase process. We first screened titles and abstracts for potential relevance. Then, we retrieved all records classified as "include" or "unsure" and reviewed their full text for eligibility. Any

disagreements between reviewers were resolved by discussion or third-reviewer consultation when necessary.

#### **Data extraction**

Reviewers used a standardized form to extract data in Microsoft Office Excel (v. 2016, Microsoft Corporation, Redmond, WA). One reviewer independently extracted data from each included study and a second reviewer verified a random 10% sample. Since no major errors or omissions were noted, we did not undertake further verification.

We extracted the following data: country of publication; funding source; study design; inclusion and exclusion criteria; population characteristics (i.e., sample size, age and gender distribution, physician specialty); setting (i.e., physician workplace, urban or rural); exposure or intervention; definition of fatigue or insufficient sleep; sleep and fatigue scales used and timing of measurement; comparators (if applicable); and outcomes.

# Risk of bias appraisal

Two reviewers independently assessed the risk of bias in each included study using standard tools. Disagreements were resolved via discussion or by consulting a third reviewer. We used the Cochrane Risk of Bias tool[22] to assess randomised controlled trials. Adapted versions of the tool developed by the Effective Practice and Organization of Care group[26] were used to assess before-after and time series studies. We used the Newcastle-Ottawa Quality Assessment Scale[27] to appraise cohort studies. We adapted the scale to assess cross-sectional studies and the one non-comparative study.

#### **Evidence synthesis**

We considered clinical and methodological heterogeneity in our decision on whether to proceed with meta-analysis for the outcomes identified. For most outcomes, we found high levels of heterogeneity in study design, populations, exposures or interventions, and outcome measures and chose not pool the data via meta-analysis. Thus, we have presented the findings for most outcomes narratively and in summary tables.

When statistical pooling was appropriate, this was undertaken using Review Manager (RevMan v.5.3, Copenhagen: The Nordic Cochrane Centre, the Cochrane Collaboration, 2014) via pairwise meta-analysis

When data were not presented in the format required for meta-analysis, we estimated means or standard deviations (SDs) using standard equations. We used the median instead of the mean for one study[31] for the outcomes of length of stay and operating time. Additionally, for one study[32] in the length of stay analysis where the SD could not be estimated, we substituted the mean variance of other studies within the meta-analysis.[33]

# **RESULTS**

We identified 16,083 unique records via the database searches, 56 grey literature sources, and 14 additional records in reference lists of systematic reviews. We excluded 15,016 citations by title and abstract, and another 1,090 by full text. Forty-seven studies[31, 32, 34-78] were eligible for inclusion, and 6[31, 32, 41, 58, 63, 77] were included in meta-analysis for the outcomes of operating time, intra-and post-operative complications, patient mortality and length of hospital stay. Figure 1 shows the flow of studies through the selection process.

### **Included study characteristics**

A summary of the study characteristics is provided in Table 1. Supplementary file 2 presents descriptive information for each included study. There were 45 observational studies [31, 32, 34-39, 41-74, 76-78] and two intervention studies. [40, 75] All studies were quantitative. Nearly half (n = 20/47, 43%) of the studies took place in North America, [31, 32, 35, 37, 38, 41, 45, 48, 57-60, 62, 63, 65-68, 72, 77] and

Table 1. Summary characteristics of the included studies

Study characteristics	n	%	Physician characteristics	n	%	Exposures, interventions and	n	%
						outcomes		
Study design			Gender			Exposures (observational) <sup>a</sup>	45	96
Cross-sectional	34	72	Reported <sup>b</sup>	38	81	Fatigue-related	15	32
Cohort	6	13	>50% male	30	79	Sleep-related	37	79
Before-after	3	6	Age		•	Overnight or extended shifts	18	38
RCT	2	4	Reported <sup>b</sup>	38	81	Interventions (experimental)	2	4
Time series	1	2	Range (years)	20 to	>70	Outcomes		
Non-comparative	1	2	Specialty area <sup>c</sup>			Physician health and wellbeing	28	60
Region and country	•		Surgeons	13	28	Work and life satisfaction	9	19
North America	20	43	Anesthesiologists	10	21	Burnout	7	15
US	15	32	Generalists	7	15	Stress	8	17
Canada	4	9	ED or ICU physicians	3	6	Mental health and wellbeing	7	15
Canada, US & Mexico	1	2	Oncologists	2	4	Other health-related outcomes	5	11
Europe	16	34	Obstetrician-gynecologists	1	2	Physician performance, risk of error	21	45
France	4	9	Mixed groups	14	30	Psychomotor performance	7	15
Finland	3	6	Work setting <sup>d</sup>			Work ability and quality of care	5	11
Spain	2	4	Hospitals	37	79	Incidence of medical errors	5	11
Austria	2	4	Private practice	13	28	Surgical efficiency, effectiveness	6	13
Norway	2	4	Primary care centres, outpatient clinics	7	15	Patient outcomes	6	13
Denmark	1	2	Academic practice, training programs	5	11	1//,		-
Germany	1	2	Other (e.g., industry, military)	11	23			
Malta	1	2	Not reported	3	6			
Japan	4	9	Urban or rural					
Australia	2	4	Reported <sup>b</sup>	16	34			
Israel	2	4	Urban	12	75			
New Zealand	2	4	Rural	2	13			
United Kingdom	1	2	Mixed	2	13			

ED: emergency department; ICU: intensive care unit; RCT: randomised controlled trial; US: United States of America

<sup>&</sup>lt;sup>a</sup>Exposures that have been directly related to an outcome. Some studies included multiple exposures.

<sup>b</sup>Percentages presented using the total number of studies where the outcome was reported as the denominator.

Anesthesiologists include physician anesthetists; generalists include primary care physicians, internists, and general practitioners; mixed groups refers to studies including more than one physician group or specialty (usually large-scale surveys). In some studies, multiple distinct groups were represented.

<sup>d</sup>As defined by the authors. Values for the settings will exceed 100% because studies may occur in more than one setting.



Fifteen (32%) studies reported on fatigue-related exposures (e.g., as a source of stress, exhaustion, physical fatigue; hereafter referred to as 'fatigue'),[35, 40, 45, 48, 57, 63-71, 73] while others (n = 37/47, 79%) reported on sleep-related exposures (e.g., sleep hours, insufficient sleep, sleep deprivation, sleep disruption, sleepiness; hereafter referred to as 'insufficient sleep').[31, 32, 34, 36-47, 49-56, 58-62, 64, 67, 71, 72, 74-78] A few (n = 5/47, 11%) reported on both.[40, 45, 64, 67, 71] In some cases (n = 18/47, 38%), fatigue or insufficient sleep were related to overnight work or long on-call shifts.[31, 32, 34, 37, 40, 41, 43, 45, 46, 50, 51, 53, 58-60, 63, 75, 77] Measured outcomes varied widely and were ultimately organised into physician physical and mental health, physician performance and risk of error, and patient outcomes.

#### Risk of bias appraisal

The overall quality of the body of research was poor; 62% (n = 29/47) of studies were rated at unclear or high risk of bias. Of the two randomised controlled trials, one was rated as unclear overall risk of bias[75] and one as high risk.[40] All cohort studies were at low risk of bias (mean score: 8.4/9, range: 8-9).[31, 32, 41, 58, 63, 77] All of the before-after studies were rated as high risk of bias.[34, 45, 50] The single time series study was assessed at high risk of bias.[51] The cross-sectional studies varied in performance (mean score: 3.0/5, range: 1-4); only one-third (n = 12/34, 35%) were at low risk of bias.[39, 42, 44, 47, 52, 59, 60, 69, 70, 72, 74, 76] The one non-comparative study was at unclear risk of bias.[43] Detailed assessments of the sources of bias per study are shown in Supplementary file 3.

# Physician health and wellbeing outcomes

Twenty-eight studies reported on physician health and wellbeing-related outcomes, [35, 36, 38-40, 42, 46-48, 50-57, 60, 62, 64, 67, 68, 70-72, 74, 76, 78] including burnout (n = 7), stress (n = 8), mental health and wellbeing (n = 7), life and job satisfaction (n = 9) and other markers of health (n = 5) (Supplementary file 4).

Seven cross-sectional studies reported on burnout (5 low[39, 70, 72, 74, 76], 1 unclear[54], 1 high risk of bias[62]) among surgeons,[54, 62] anesthesiologists,[39] generalists,[76] and other mixed groups.[70, 72, 74] Two studies reported on surgeons; the larger (n = 2,564, low risk of bias) study of neurosurgeons showed increased odds of burnout with sleep deprivation (hours of sleep per night; OR 0.84, 95% CI 0.75 to 0.94, P = 0.002).[54] Among anesthesiologists one study (n = 565, low risk of bias) indicated that burnout (measured via Maslach Burnout Inventory) was more prevalent among the sleep-deprived ('lack of sleep' on one question; 47.6% vs. 16.3%, P < 0.001).[39] In one small (n = 11) study of generalists, those with burnout (measured via Pines Burnout Measure) had poorer Pittsburgh Sleep Quality Index scores (7.24±4.17 vs. 2.72±2.22, P < 0.001).[76] In the two larger studies of mixed physician groups (low risk of bias), burnout (measured via 5-point scale) was more prevalent among those who were sleep deprived (<7 hours of sleep per 24 hours; 39.6% vs. 26.4%, P < 0.05),[72] and physical fatigue ('feeling tired' on a 7-point scale) was correlated with burnout (Shirom-Melamed Burnout Measure; r = 0.88, P < 0.05).[70] In summary, evidence from 7 cross-sectional studies (71% at low risk of bias), showed associations between insufficient sleep and burnout.

Six cross-sectional studies (2 low[47, 52], 1 unclear[46], 3 high risk of bias[35, 62, 64]), one uncontrolled before-after study (high risk of bias[50]), and one intervention study (high risk of bias[40]) reported on stress outcomes among surgeons,[62] anesthesiologists,[50, 52] emergency physicians,[40, 64] internal medicine physicians,[46] and mixed groups.[35, 47] In a small sample (n = 20) of internal medicine physicians, insufficient sleep related to a 24-hour call shift showed no association with biochemical or physiological stress parameters, except levels of thyroid stimulating hormone, which was higher post-shift (P = 0.049, data not reported).[46] The remaining observational studies suggested that there was an association between insufficient sleep or fatigue and stress. The one study of orthopedic surgeons (n = 264, high risk of bias) showed that insufficient sleep (measured on a 3-point scale) and psychological distress (measured via General Health Questionnaire-12) were correlated (data not reported, P < 0.001).[62] The two reports on anesthesiologists were of varied quality; the larger (n = 328, low risk of

Seven cross-sectional studies (2 low, [52, 60] 3 unclear, [67, 71, 78] 2 high risk of bias [36, 53]) reported on aspects of mental health including addiction or substance misuse, [36, 53, 71] depression, [78] thoughts of suicide,[52] mood disturbance,[60, 71] and overall wellbeing.[67] One study,[53] which was at high risk of bias, showed no association between hours of sleep when on call and hazardous drinking behaviours (via Alcohol Use Disorder Identification Test). Meanwhile, the six other studies all showed associations between insufficient sleep and fatigue and reduced mental health. Three studies reported on anesthetists, [36, 52, 60] with two large surveys showing increased odds of tobacco (OR 1.42, 95% CI 1.04 to 1.94) and tranquilizer/hypnotics (OR 3.26, 95% CI 2.12 to 5.02) dependency being predicted by sleep deprivation (measured by one question),[36] and sleep disturbance being associated with thoughts of suicide (using a 4-point scale; P = 0.009).[52] A small study (n = 21) showed greater mood disturbance following a 17-hour night shift than a usual day (Profile of Mood States score 42.57±15.26 vs. 70.90±6.91, P < 0.001).[60] Among oncologists (n = 241), overall wellbeing was predicted by lower levels of fatigue after controlling for personal and professional characteristics (assessed via linear analog scale quality of life survey, P = 0.002).[67] A large (n = 3,862, unclear risk of bias) study of physicians showed that insufficient sleep (lower sleep hours when not at work in the past month) was associated with increased odds of depression (Quick Inventory Depressive Scale; OR 2.70, 95% CI 1.82 to 4.03 for men; OR 2.38, 95% CI 1.11 to 5.10 for women).[78] In open-ended questions, senior physicians in one study (unclear risk of bias) attributed the development of mental illness to tiredness and stress at work.[71] In summary, 7 cross-sectional studies (29% at low risk of bias) were identified, and of these 6 supported an association between insufficient sleep or fatigue and negative mental health outcomes.

Nine cross-sectional studies (4 low,[42, 47, 72, 74] 2 unclear,[55, 68] 3 high risk of bias[38, 48, 62]) reported on outcomes related to job satisfaction,[42, 47, 48, 55, 72, 74] life satisfaction,[38, 62, 72] or work-life balance.[68, 72] The six studies that investigated job satisfaction were all at low risk of bias and generally included mixed groups of physicians;[47, 72, 74] one study reported on general practitioners,[55], another on surgeons,[48] and one on mixed specialties.[42] Three studies showed that reductions in sleep duration and/or quality[47, 48, 74] were associated with reduced job satisfaction. Meanwhile one showed no association between insufficient sleep (<7 hours per 24-hour period) and career satisfaction (measured on a 5-point Likert scale),[72] and another showed no relationship between earlier sleep disturbance (Jenkins Scale) and later job demands or job control (measured via 5-point scale).[42] A single study (n = 92) reporting on rural general practitioners indicated that frequent sleep disturbance (measured on a 7-point scale) predicted the intention to retire early (OR 2.91, 95% CI 1.11 to 7.6, P < 0.05).[55] In summary, 6 cross-sectional studies (all at low risk of bias) were identified, and all but two[42, 72] of these studies showed that insufficient sleep and fatigue were associated with reductions in satisfaction.

The three studies reported on life satisfaction.[38, 62, 72] Of two studies among mixed physician groups, [38, 72] the one larger (n = 840) study showed that insufficient sleep (< 7 hours per day) was a predictor of reduced life satisfaction (measured on a 5-point Likert scale; OR 0.44, 95% CI 0.29 to 0.67, P ≤ 0.05).[72] One study at high risk of bias reported on orthopedic surgeons (n = 264), showing that sleep deprivation (measured via 3-point scale) was correlated with lower marital satisfaction (Revised Dyadic Adjustment Scale; data not reported, P < 0.001).[62] Two large studies at low or unclear risk of bias reported on work-life balance.[68, 72] Among oncologists (n = 1,117), reduced satisfaction with work-life balance (measured on a 5-point Likert scale) was predicted by high levels of fatigue (measured via 10-point visual analog scale), even when controlling for personal and work-related factors and burnout (OR 0.489, 95% CI 0.337 to 0.710, P < 0.001).[68] Among a mixed group of physicians (n = 840, low risk of bias), insufficient sleep (<7 hours in a typical 24-hour period) predicted a reduced perception of having balanced personal and professional commitments (5-point Likert scale; OR 0.46, 95% CI 0.31 to 0.71, P ≤ 0.05).[72] In summary, 3 cross-sectional studies (all unclear or high risk of bias) supported an association between insufficient sleep or fatigue and reduced life satisfaction, and 2 cross-sectional studies (50% low risk of bias) supported an association with reduced work-life balance.

Four cross sectional studies (3 unclear, [56, 57, 71] 1 high risk of bias [38]) and one time series study (high risk of bias [51]) reported on other health-related outcomes. Among a mixed group of physicians (n = 180), one study at high risk of bias showed that Epworth Sleepiness Scale scores were higher among physicians who worried about having a car accident while driving home (7.0 vs. 5.4, P < 0.001).[38] Among generalists (n = 578), almost 1 in 10 (8.7%) admitted to falling asleep while driving due to fatigue.[57] Also among generalists (n = 92), those with frequent work-related sleep disturbance (measured on a 7-point scale) were at increased odds of sickness presenteeism (OR 2.92, 95% CI 1.19 to 7.16, P = 0.02).[56] The one time series study concluded that a single 24-h shift did not cause major chronodisruption (based on serum melatonin measurement) among anesthetists (n = 10).[51] Meanwhile, open-ended comments from a large sample (n = 3,550) of senior physicians suggests that they attributed the development of physical health problems to a lifestyle of insufficient sleep, poor eating habits and lack of exercise imposed by their jobs.[71] In summary, 5 cross sectional studies (0% at low risk of bias) supported associations between insufficient sleep and fatigue and varied deleterious health outcomes (i.e., car accidents, sickness presenteeism, physical health problems). One time series study at high risk of bias did not support such a relationship.

# Physician performance and risk of errors

Twenty-one studies reported on physician performance and safety-related outcomes, [31, 32, 34, 37, 38, 41, 43-47, 49, 50, 59, 61, 65, 66, 69, 71, 73, 75] including surgical efficiency and effectiveness (n = 6), psychomotor performance (n = 7), work ability and quality of care (n = 5) and medical errors (n = 5) (Supplementary file 4).

Four cohort studies (all low risk of bias[31, 32, 41, 63]), one before-after study (high risk of bias[34]) and one randomized controlled trial (high risk of bias[75]) examined the effects of insufficient sleep from overnight work or extended shifts, during surgeries[31, 32, 41] or laparoscopic simulations.[34, 75] We pooled the data from these studies[31, 32, 41, 63] via meta-analysis, which showed no difference in operating time (sometimes referred to as surgeon efficiency) between sleep deprived and non-sleep deprived surgeons (Figure 2; n = 50,046, MD -0.14, 95% CI -1.60 to 1.33,  $I^2 = 0\%$  (P = 0.70)). Of studies not meta-analysed, the small (n = 29) before-after study showed no impact of sleep deprivation from shift-work nor of sleep hours on performance on a laparoscopic simulation (LapSimGyn).[34] One small (n = 64) intervention study compared a 24-hour shift to a usual work day, also finding no detriment to performance on a laparoscopic simulation (Minimally Invasive Surgical Trainer-Virtual Reality) despite

diminished sleep hours while working on-call.[75] In summary, pooled data from 4 cohort studies (100% low risk of bias) showed no effect of insufficient sleep on surgical efficiency. Additional data from one RCT (high risk of bias) and one before-after study (high risk of bias) also showed no association between insufficient sleep and performance on laparoscopic simulations.

Two before-after studies (high risk of bias[45, 50]) and five cross-sectional studies (2 low,[43, 59] 3 unclear, [37, 46] 1 high risk of bias [61]) reported on psychomotor performance outcomes among surgeons,[45] anesthesiologists,[37, 43, 50, 59] emergency physicians,[61] and internal medicine physicians.[46] Among a small group of surgeons (n = 9), performance on a virtual ring transfer task deteriorated after an on-call shift (data not reported, P < 0.05).[45] The four studies among anesthetists reported mixed findings. One small (n = 11) before-after study showed longer reaction times (690.8±73.4 vs. 746.5±113.7 milliseconds) and reduced concentration ability (26.4±23.5 vs. 56.3±23.0 on a 100-point scale, P = 0.007) following a 24-hour shift with insufficient sleep; [50] Two others found that insufficient sleep due to overnight shifts was associated with slower reaction times.[43, 59] Conversely, a small study (n = 11) found no effect of overnight shiftwork with insufficient sleep on any measure of psychomotor performance except Hopkin's Verbal Learning Test (t-score of 48.6±7.6 vs. 41.5±9.9, P = 0.04).[37] Among emergency physicians (n = 18), one study (high risk of bias) showed that those who were sleep deprived (<5 hours sleep after a 24-hour shift) had a reduced performance on most but not all psychomotor tests (Battery Test Reaction 5),[61] while among internal medicine physicians (n = 20, low risk of bias), neurocognitive parameters did not seem to worsen post-call.[46] In summary, two before-after (0% low risk of bias) and 5 cross-sectional studies (40% low risk of bias) showed mixed results for the association between fatigue or insufficient sleep and psychomotor performance.

Five cross-sectional studies (2 low,[47, 69] 1 unclear,[71] 2 high risk of bias[38, 65]) reported on associations between sleep deprivation or fatigue and work ability or perceived performance, all among mixed groups of physicians.[38, 47, 65, 69, 71] The two large studies at low risk of bias showed that sleep problems and fatigue were inversely associated with physicians' perceived quality of work.[47, 69] Among 1,541 physicians in Finland, sleeping problems (measured by 4 questions from the Jenkins Scale) were inversely associated with scores on the Work Ability Index ( $\beta$  = -0.29, P < 0.001),[47] while a study of 890 physicians from Israel demonstrated that perceived quality of care was predicted by fatigue (1 item on the Shirom-Melamed Burnout Measure) even after controlling for components of burnout ( $\beta$  = 0.17, P < 0.05).[69] Similarly, in one study, comments from senior physicians suggested that continual

tiredness and exhaustion negatively affected their perceived competence. [71] The two studies [38, 65] that were at high risk of bias had conflicting findings. In summary, 5 cross-sectional studies (40% at low risk of bias) reported on perceived work performance; those that were at low risk of bias supported an association between fatigue or insufficient sleep and reduced performance.

Five cross-sectional studies (1 low,[44] 2 unclear,[49, 66] 2 high risk of bias[38, 73]) reported on associations between insufficient sleep or fatigue and self-reported medical errors among surgeons,[66] anesthesiologists[44] and mixed groups of physicians.[38, 49, 73] A large (n = 7,905) study at unclear risk of bias showed that only 6.9% of surgeons reported fatigue as the most important contributor to medical errors.[66] Among anesthesiologists, a smaller study (n = 183) at low risk of bias showed that the risk of self-reported fatigue-related errors increased with more nights of work-related sleep disturbance (RR 1.25, 95% CI 1.06 to 1.49).[44] Two of the studies reporting on mixed groups of physicians had conflicting results,[38, 49] while another reported that physicians' opinions on the association between fatigue and prescribing errors differed by work setting.[73] One-third (34%) of community-based, 96% of hospital-based, and 8% of office-based physicians believed that there was a high or very high association between fatigue and prescribing errors (P < 0.05).[73] In summary, 5 cross-sectional studies (20% at low risk of bias) reported on self-reported errors, and these showed mixed findings for associations with fatigue or insufficient sleep.

#### **Patient Outcomes**

Six large (n = 270 to 38,978) cohort studies at low risk of bias reported on patient outcomes, all related to surgical[31, 32, 41, 58, 63, 77] or obstetric[58] procedures (Supplementary file 4). In these studies, insufficient sleep or fatigue were typically defined as overnight work prior to a daytime procedure[31, 41, 58, 63, 77]; though two studies measured sleep hours[32] or 'sleep opportunity'.[58] We pooled data for procedures performed by sleep deprived versus non-sleep deprived surgeons (or obstetrician-gynecologists in one case[58]). Analyses showed no difference in the rate of post-operative complications (Figure 3; 5 studies,[31, 32, 41, 63, 77] n = 60,201, RR 0.99, 95% Cl 0.95 to 1.03,  $I^2 = 0\%$  (P = 0.45) nor patient mortality (Figure 4; 5 studies,[31, 32, 41, 63, 77] n = 60,436, RR 0.98, 95% Cl 0.84 to 1.15,  $I^2 = 0\%$  (P = 0.73)). One study[77] in the mortality analysis reported the number of deaths only as  $\leq$ 5. We assumed 2 events for this study (midpoint between 0 and 5); sensitivity analysis using the lowest (i.e., 0) and highest (i.e., 5) possible number of events did not change the overall result (Supplementary file 5). We found considerable between-study heterogeneity in the analyses for intraoperative

complications ( $I^2 = 82\%$ ) and length of stay ( $I^2 = 86\%$ ), which could not be explained via subgroup analyses by procedure type, thus we have suppressed the average estimates of effect. For length of stay, the results of one study on cardiac surgeries favoured sleep deprived surgeons,[32] while the others[31,41,63] had null results. For intraoperative complications, the findings of one study[63] favoured non-sleep deprived surgeons, but the others[58,77] had null results.

# **DISCUSSION**

Fatigue and chronic insufficient sleep are two potential drivers of reduced physician wellbeing[17, 19] that have thus far been understudied in physicians in independent practice. Burnout is becoming increasingly prevalent among physicians,[14-16] and recent research indicates that comprehensive individual- and system-level strategies are needed to address the problem.[6-9, 19, 21] We have systematically reviewed evidence from a heterogeneous array of available studies reporting on diverse outcomes related to physicians in independent practice and their patients. The included studies were often at high or unclear risk of bias, included small samples of physicians, and inconsistently measured and reported exposures and outcomes. The key message gleaned from this review is that despite growing interest in the topic of physician wellness, the robust evidence needed to inform individual and systems-level fatigue management strategies is lacking.

Traditionally, much of the fatigue-related research has focused on hazards to patients. The current review included six cohort studies showing that insufficient sleep and/or fatigue did not seem to result in increased rates of patient morality or post-operative complications; findings for length of stay and intra-operative complications were inconclusive. Evidence for psychomotor performance, surgical skills and errors suggest that there is indeed a potential for negative outcomes. The included studies, like many of the others in this and other systematic reviews,[79] employed indirect definitions that make it difficult to classify sleep deprived physicians with certainty. In recent years there has been a shift away from the singular focus on patient safety toward a more comprehensive view that also considers the detrimental effects of fatigue, sleep loss and other occupational hazards on physician wellness.[80] Evidence from this review supports that fatigue and insufficient sleep may be negatively associated with physician health and wellbeing. It is now recognized that health systems cannot be sustained by a workforce that is facing an epidemic of burnout.[19, 81, 82]

The most salient finding of this review is that the current evidence is insufficient to inform policy and practice. Correspondingly, a 2016 research summit on physician wellness and burnout outlined the need for timely, relevant and methodologically robust research to inform practice and policy.[21] The findings herein may be used as motivation for researchers and practitioners to develop and design methodologically strong research programs related to physician fatigue, inform successful research grant proposals, and lobby healthcare organizations to increase the focus on physician fatigue management programs. It will be important to make use of existing validated measures[87-89] consistently in future research. Identifying outcomes of importance to physicians and their patients should be prioritized, such that these may be collected within intervention studies. Reporting these consistently will allow for the effective synthesis of findings and reduce research waste.[90] Integrated knowledge translation strategies involving multiple stakeholder groups (e.g., physicians, patients, medical schools, physicians' associations and governing bodies, policymakers) may help to ensure that the research is relevant and facilitates decision-making.[91]

# **Strengths and Limitations**

Our systematic review is the first to synthesize evidence on the effects of fatigue and insufficient sleep on physicians in independent practice. The review is timely, given recent calls for research into individual and organisational solutions for burnout, [20, 21] and an increased focus on physician

health.[80, 81] While we have identified a diverse body of evidence, we could not draw definitive conclusions due to methodological weaknesses (e.g., 62% at high risk of bias, reliance primarily on cross-sectional designs and uncontrolled studies, subjective measurement of exposures and outcomes, small sample sizes, inclusion of predominantly male physicians within urban settings) and heterogeneous outcome measures in the included studies. Given that the 2017 update search was limited to one database, it is possible that a small number of relevant studies could have been missed. We believe that the likelihood that these might alter the conclusions of the review is low. The findings may have been influenced by publication bias, and may not be generalized to all settings, given our restriction to high income countries. Confidence in the conclusions is limited due to multiple comparisons.

#### CONCLUSION

The evidence synthesized in this review suggests that fatigue and insufficient sleep are associated with some detrimental physician health and wellbeing outcomes; the evidence for potential associations with performance and safety outcomes was mixed. Meta-analyses for patient outcomes demonstrated that in many cases, potential relationships with physician sleep deprivation remain unclear. Our overall confidence in the findings is low, owing to multiple comparisons and a body of research that is hindered by methodological weaknesses. Further methodologically robust research that includes consistent outcomes that are of interest to physicians and their patients is needed to inform strong practice recommendations and policy decisions.

## **ACKNOWLEDGMENTS**

We would like to thank Dr. Meghan Sebastianski and Andrea Milne for assistance with screening, and Dr. Allison Gates for screening, data verification, and quality assessment of the French studies.

# **COPYRIGHT/LICENSE FOR PUBLICATION**

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide license to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

# **DATA ACCESS STATEMENT**

All authors, external and internal, had full access to all of the data in the study and can take responsibility for the integrity of the data and the accuracy of the interpretation.

#### **DATA SHARING STATEMENT**

The data pertaining to this systematic review are available from the corresponding author upon reasonable request.

### **FIGURE CAPTIONS**

- Figure 1. Flow of records through the selection process
- Figure 2. Forest plot for operating time among sleep deprived and non-sleep deprived surgeons
- **Figure 3.** Forest plot for post-operative complications among surgeries performed by sleep deprived and non-sleep deprived surgeons
- **Legend:** Vinden 2013 reported iatrogenic injuries; Schieman 2008, Govindarajan 2015, and Chu 2011 reported post-operative complication rate; Ellman 2004 reported post-operative complications (other types of complications reported not included in the analysis)
- **Figure 4.** Forest plot for patient mortality among surgeries performed by sleep deprived and non-sleep deprived surgeons

- 1. Greig P, Snow R. Fatigue and risk: are train drivers safer than doctors? *BMJ* 2017;359. doi: https://doi.org/10.1136/bmj.i5107
- 2. Parker JBR. The effects of fatigue on physician performance an underestimated cause of physician impairment and increased patient risk. *Can J Anaesth* 1987;34(5):489-95.
- 3. Temple J. Resident duty hours around the globe: where are we now? BMC Med Educ 2014;14(Suppl 1):S8. https://doi.org/10.1186/1472-6920-14-S1-S8
- 4. Imrie K, Frank J, The National Steering Committee on Resient Duty Hours. Fatigue, risk, & excellence: towards a pan-Canadian consensus on resident duty hours. Ottawa, Ontario: The Royal College of Physicians and Surgeons of Canada, 2013. <a href="http://www.residentdutyhours.ca/final\_report.php">http://www.residentdutyhours.ca/final\_report.php</a> (accessed 12 Jan 2018).
- Accreditation Council for Graduate Medical Education. History of duty hours. 2017.
   <a href="http://www.acgme.org/What-We-Do/Accreditation/Clinical-Experience-and-Education-formerly-Duty-Hours/History-of-Duty-Hours">http://www.acgme.org/What-We-Do/Accreditation/Clinical-Experience-and-Education-formerly-Duty-Hours/History-of-Duty-Hours</a> (accessed 12 Jan 2018).
- 6. Bolster L, Rourke L. The effect of restricting residents' duty hours on patient safety, resident well-being, and resident education: an updated systematic review. *J Grad Med Educ* 2015;7(3):349-63.
- 7. Harris JD, Staheli G, LeClere L, Andersone D, McCormick F. What effects have resident work-hour changes had on education, quality of life, and safety? A systematic review. *Clin Orthop* 2015;473(5):1600-8.
- 8. Ahmed N, Devitt KS, Keshet I, et al. A systematic review of the effects of resident duty hour restrictions in surgery: impact on resident wellness, training, and patient outcomes. *Ann Surg* 2014;259(6):1041-53.
- 9. Peets A, Ayas NT. Restricting resident work hours: the good, the bad, and the ugly. *Crit Care Med* 2012;40(3):960-6.
- Canadian Medical Association. CMA Policy: Management of physician fatigue. Ottawa, Canada:
   Canadian Medical Association; 2014.
- 11. Anim M, Markert RJ, Wood VC, Schuster BL. Physician practice patterns resemble ACGME duty hours. *Am J Med* 2009;122:587-93.
- 12. Transport Canada. Fatigue risk management system for the canadian aviation industry: policies and procedures development guidelines. Ottawa, Canada: Her Majesty the Queen in Right of Canada, as

represented by the Minister of Transport, 2008. https://www.tc.gc.ca/eng/civilaviation/publications/TP14576-6042.htm (accessed 12 Jan 2018).

- 13. Patterson PD, Higgins JS, Lang ES, et al. Evidence-based guidelines for fatigue risk management in ems: formulating research questions and selecting outcomes. *Prehosp Emerg Care* 2017;21(2):149-56.
- 14. Boudreau RA, Grieco RL, Cahoon SL, Robertson RC, Wedel RJ. The pandemic from within: two surveys of physician burnout in Canada. *Can J Commun Ment Health* 2006;25(2):71-88.
- 15. Kumar S. Burnout and doctors: prevalence, prevention and intervention. *Healthcare* 2016;4(3):37. doi: 10.3390/healthcare4030037
- 16. Shanafelt TD, Boone S, Tan L, et al. Burnout and satisfaction with work-life balance among us physicians relative to the general us population. *Arch Intern Med* 2012;172(18):1377-85.
- 17. Brady KJS, Trockel MT, Khan CT, et al. What do we mean by physician wellness? A systematic review of its definition and measurement. *Acad Psychiatry* 2017. doi: 10.1007/s40596-017-0781-6
- 18. Fralick M, Flegel K. Physician burnout: who will protect us from ourselves? *CMAJ* 2014;186(10):731. doi: 10.1503/cmaj.140588
- 19. Shanafelt TD, Noseworthy JH. Executive leadership and physician well-being: nine organizational strategies to promote engagement and reduce burnout. *Mayo Clin Proc* 2017;92(1):129-46.
- 20. Lemaire JB, Wallace JE. Burnout among doctors. BMJ 2017;358:j3360. doi: 10.1136/bmj.j3360
- 21. Dyrbye LN, Trockel M, Frank E, et al. Development of a research agenda to identify evidence-based strategies to improve physician wellness and reduce burnout. *Ann Intern Med* 2017;166(10):743-4.
- 22. Higgins JPT, Green S (editors). The Cochrane handbook for systematic reviews of interventions, version 5.1.0. London, UK: The Cochrane Collaboration, 2011. <a href="http://www.handbook.cochrane.org">http://www.handbook.cochrane.org</a> (accessed 12 Jan 2018).
- 23. 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(7):e1000100. https://doi.org/10.1371/journal.pmed.1000100
- 24. Pattani R, Wu PE, Dhalla IA. Resident duty hours in Canada: past, present and future. *CMAJ* 2014;186(10):761-5.
- 25. International Monetary Fund. World economic outlook: too slow for too long. Washington, DC: International Monetary Fund, 2016. <a href="http://www.imf.org/external/pubs/ft/weo/2016/01/">http://www.imf.org/external/pubs/ft/weo/2016/01/</a> (accessed 12 Jan 2018).

- 27. Wells GA SB, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa, Canada: The Ottawa Hospital Research Institute, 2014.
  http://www.ohri.ca/programs/clinical\_epidemiology/oxford.asp (accessed 12 Jan 2018).
- 28. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-88.
- 29. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-60.
- 30. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315(7109):629-34.
- 31. Govindarajan A, Urbach DR, Kumar M, et al. Outcomes of daytime procedures performed by attending surgeons after night work. *N Engl J Med* 2015;373(9):845-53.
- 32. Chu MW, Stitt LW, Fox SA, Kiaii B, Quantz M, Guo L, et al. Prospective evaluation of consultant surgeon sleep deprivation and outcomes in more than 4000 consecutive cardiac surgical procedures. *Arch Surg* 2011;146(9):1080-5.
- 33. Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *J Clin Epidemiol* 59(1):7-10.
- 34. Amirian I, Andersen LT, Rosenberg J, Gogenur I. Laparoscopic skills and cognitive function are not affected in surgeons during a night shift. *J Surg Educ* 2014;71(4):543-50.
- 35. Aziz A. Sources of perceived stress among american medical doctors: a cross-cultural perspective. Cross Cultural Management 2004;11(4):28-39.
- 36. Beaujouan L, Czernichow S, Pourriat JL, Bonnet F. Prevalence and risk factors for substance abuse and dependence among anaesthetists: a national survey. *Ann Fr Anesth Réanim* 2005;24(5):471-9.
- 37. Chang LC, Mahoney JJ, 3rd, Raty SR, Ortiz J, Apodaca S, De La Garza R 2nd. Neurocognitive effects following an overnight call shift on faculty anesthesiologists. *Acta Anaesthesiol Scand* 2013;57(8):1051-7.
- 38. Chen I, Vorona R, Chiu R, Ware JC. A survey of subjective sleepiness and consequences in attending physicians. *Behav Sleep Med* 2008;6(1):1-15.

- 39. Doppia MA, Estryn-Béhar M, Fry C, Guetarni K, Lieutaud T, comité de pilotage de l'enquête SESMAT. Burnout in French doctors: a comparative study among anaesthesiologists and other specialists in French hospitals (SESMAT study). *Ann Fr Anesth Réanim* 2011;30(11):782-94.
- 40. Dutheil F, Trousselard M, Perrier C, et al. Urinary interleukin-8 is a biomarker of stress in emergency physicians, especially with advancing age--the JOBSTRESS randomized trial. *PLoS One* 2013;8(8):e71658. https://doi.org/10.1371/journal.pone.0071658
- 41. Ellman PI, Law MG, Tache-Leon C, et al. Sleep deprivation does not affect operative results in cardiac surgery. *Ann Thorac Surg* 2004;78(3):906-11.
- 42. Elovainio M, Heponiemi T, Jokela M, et al. Stressful work environment and wellbeing: what comes first? *J Occup Health Psychol* 2015;20(3):289-300.
- 43. Gander P, Millar M, Webster C, Merry A. Sleep loss and performance of anaesthesia trainees and specialists. *Chronobiol Int* 2008;25(6):1077-91.
- 44. Gander PH, Merry A, Millar MM, Weller J. Hours of work and fatigue-related error: a survey of New Zealand anaesthetists. *Anaesth Intensive Care* 2000;28(2):178-83.
- 45. 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(6):813-9.
- 46. Harbeck B, Suefke S, Haas CS, Lehnert H, Kropp P, Moenig H. No stress after 24-hour on-call shifts? *J Occup Health* 2015;57(5):438-47.
- 47. Heponiemi T, Puttonen S, Elovainio M. On-call work and physicians' well-being: testing the potential mediators. *Occup Med* 2014;64(5):352-7.
- 48. Jackson TN, Pearcy CP, Khorgami Z, Agrawal V, Taubman KE, Truitt MS. The physician attrition crisis: a cross-sectional survey of the risk factors for reduced job satisfaction among US surgeons. *World J Surg* 2017;24. doi: 10.1007/s00268-017-4286-y
- 49. Kaneita Y, Ohida T. Association of current work and sleep situations with excessive daytime sleepiness and medical incidents among Japanese physicians. *J Clin Sleep Med* 2011;7(5):512-22.
- 50. Lederer W, Kopp M, Hahn O, et al. Post-duty psychomotor performance in young and senior anaesthetists. *Eur J Anaesthesiol* 2006;23(3):251-6.
- 51. Leichtfried V, Putzer G, Perkhofer D, Schobersberger W, Benzer A. Circadian melatonin profiles during single 24-h shifts in anesthetists. *Sleep Breath* 2011;15(3):503-12.
- 52. Lindfors PM, Nurmi KE, Meretoja OA, et al. On-call stress among Finnish anaesthetists. *Anaesthesia* 2006;61(9):856-66.

- 54. Nishimura K, Nakamura F, Takegami M, et al. Cross-sectional survey of workload and burnout among Japanese physicians working in stroke care: the nationwide survey of acute stroke care capacity for proper designation of comprehensive stroke center in Japan (J-ASPECT) study. *Circ Cardiovasc Qual Outcomes* 2014;7(3):414-22.
- 55. Pit SW, Hansen V. Factors influencing early retirement intentions in Australian rural general practitioners. *Occup Med* 2014;64(4):297-304.
- 56. Pit SW, Hansen V. The relationship between lifestyle, occupational health, and work-related factors with presenteeism amongst general practitioners. *Arch Environ Occup Health* 2016;71(1):49-56.
- 57. Roberts DL, Shanafelt TD, Dyrbye LN, West CP. A national comparison of burnout and work-life balance among internal medicine hospitalists and outpatient general internists. *J Hosp Med* 2014;9(3):176-81.
- 58. Rothschild JM, Keohane CA, Rogers S, et al. Risks of complications by attending physicians after performing nighttime procedures. *JAMA* 2009;302(14):1565-72.
- 59. Saadat H, Bissonnette B, Tumin D, et al. Effects of partial sleep deprivation on reaction time in anesthesiologists. *Paediatr Anaesth* 2017;27(4):358-62.
- 60. Saadat H, Bissonnette B, Tumin D, et al. Time to talk about work-hour impact on anesthesiologists: The effects of sleep deprivation on Profile of Mood States and cognitive tasks. *Paediatr Anaesth* 2016;26(1):66-71.
- 61. Sanches I, Teixeira F, dos Santos JM, Ferreira AJ. Effects of acute sleep deprivation resulting from night shift work on young doctors. *Acta Med Port* 2015;28(4):457-62.
- 62. Sargent MC, Sotile W, Sotile MO, Rubash H, Barrack RL. Quality of life during orthopaedic training and academic practice. Part 1: orthopaedic surgery residents and faculty. *J Bone Joint Surg Am* 2009;91(10):2395-405.
- 63. Schieman C, MacLean AR, Buie WD, Rudmik LR, Ghali WA, Dixon E. Does surgeon fatigue influence outcomes after anterior resection for rectal cancer? *Am J Surg* 2008;195(5):684-7.
- 64. Sende J, Jbeili C, Schvahn S, Khalid M, Asaph J, Romano H, et al. Stress factors and stress consequences among emergency physicians: vational survey. *Annales Françaises de Médecine d'Urgence* 2012;2(4):224-31.

- 65. Sexton JB, Thomas EJ, Helmreich RL. Error, stress, and teamwork in medicine and aviation: Cross sectional surveys. *Hum Perform Extrem Environ* 2001;6(1):6-11.
- 66. Shanafelt TD, Balch CM, Bechamps G, et al. Burnout and medical errors among American surgeons. *Ann Surg* 2010;251(6):995-1000.
- 67. Shanafelt TD, Novotny P, Johnson ME, et al. The well-being and personal wellness promotion strategies of medical oncologists in the North Central Cancer Treatment Group. *Oncology* 2005;68(1):23-32.
- 68. Shanafelt TD, Raymond M, Kosty M, et al. Satisfaction with work-life balance and the career and retirement plans of US oncologists. *J Clin Oncol* 2014;32(11):1127-35.
- 69. Shirom A, Nirel N, Vinokur AD. Overload, autonomy, and burnout as predictors of physicians' quality of care. *J Occup Health Psychol* 2006;11(4):328-42.
- 70. Shirom A, Nirel N, Vinokur AD. Work hours and caseload as predictors of physician burnout: the mediating effects by perceived workload and by autonomy. *J Appl Psychol* 2010;59(4):539-65.
- 71. Smith F, Goldacre MJ, Lambert TW. Adverse effects on health and wellbeing of working as a doctor: views of the UK medical graduates of 1974 and 1977 surveyed in 2014. *J R Soc Med* 2017;110(5):198-207.
- 72. Starmer AJ, Frintner MP, Freed GL. Work-life balance, burnout, and satisfaction of early career pediatricians. *Pediatrics* 2016;e20153183. doi: 10.1542/peds.2015-3183
- 73. Tanti A, Camilleri M, Borg AA, et al. Opinions of Maltese doctors and pharmacists on medication errors. *Int J Saf Med* 2017;29(1-2):81-99.
- 74. Tokuda Y, Hayano K, Ozaki M, Bito S, Yanai H, Koizumi S. The interrelationships between working conditions, job satisfaction, burnout and mental health among hospital physicians in Japan: a path analysis. *Ind Health* 2009;47(2):166-72.
- 75. 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(6):753-7.
- 76. Vela-Bueno A, Moreno-Jimenez B, Rodriguez-Munoz A, et al. Insomnia and sleep quality among primary care physicians with low and high burnout levels. *J Psychosom Res* 2008;64(4):435-42.
- 77. Vinden C, Nash DM, Rangrej J, et al. Complications of daytime elective laparoscopic cholecystectomies performed by surgeons who operated the night before. *Obstet Gynecol Surv* 2014;69(2):71-3.

- 79. Sturm L, Dawson D, Vaughan R, et al. Effects of fatigue on surgeon performance and surgical outcomes: a systematic review. *ANZ J Surg* 2011;81(7-8):502-9.
- 80. Wallace JE, Lemaire JB, Ghali WA. Physician wellness: a missing quality indicator. *Lancet* 2009;374(9702):1714-21.

- 81. Canadian Medical Association. CMA policy: physician health. Ottawa, Canada: Canadian Medical Assocation, 2017. <a href="http://policybase.cma.ca/dbtw-wpd/Policypdf/PD18-01.pdf">http://policybase.cma.ca/dbtw-wpd/Policypdf/PD18-01.pdf</a> (accessed 12 Jan 2018).
- 82. Shanafelt T, Goh J, Sinsky C. The business case for investing in physician well-being. *JAMA Intern Med* 2017;177(2):195-205.
- 83. Gaba DM, Howard SK. Fatigue among clinicians and the safety of patients. *N Engl J Med* 2002;347(16):1249-55.
- 84. Ward S, Outram S. Medicine: in need of culture change. Itern Med J 2016;46(1):112-6.
- 85. Panagioti M, Panagopoulou E, Bower P, et al. Controlled interventions to reduce burnout in physicians: a systematic review and meta-analysis. *JAMA Intern Med* 2017;177(2):195-205.
- 86. American Medical Association. STEPSforward<sup>TM</sup>. Chicago, IL: American Medical Association, 2017. https://www.stepsforward.org/modules/joy-in-medicine (accessed 12 Jan 2018).
- 87. Jackson C. The Chalder Fatigue scale (CFQ 11). Occup Med 2015;65(1):86. doi: 10.1093/occmed/kgu168
- 88. Chalder T, Berelowitz G, Pawlikowska T, et al. Development of a fatigue scale. *J Psychosom Res* 1993;37(2):147-53.
- 89. Smets E, Garssen B, Bonke Bd, De Haes J. The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *J Psychosom Res* 1995;39(3):315-25.
- 90. Ioannidis JPA, Greenland S, Hlatky MA, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;383(9912):166-75.
- 91. Gagliardi AR, Berta W, Kothari A, Boyko J, Urquhart R. Integrated knowledge translation (IKT) in health care: a scoping review. *Implement Sci* 2016;11:38. https://doi.org/10.1186/s13012-016-0399-1

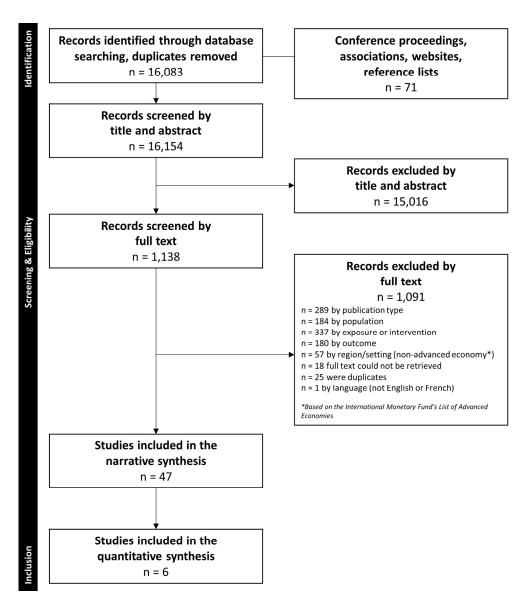


Figure 1. Flow of records through the selection process  $190x215mm (300 \times 300 DPI)$ 

BMJ Open: first published as 10.1136/bmjopen-2018-021967 on 21 September 2018. Downloaded from http://bmjopen.bmj.com/ on June 9, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

Figure 2. Forest plot for operating time among sleep deprived and non-sleep deprived surgeons

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

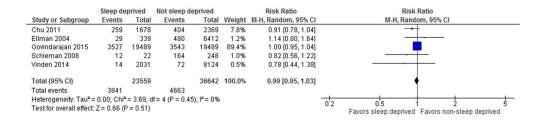


Figure 3. Forest plot for post-operative complications among surgeries performed by sleep deprived and non-sleep deprived surgeons

Legend: Vinden 2013 reported iatrogenic injuries; Schieman 2008, Govindarajan 2015, and Chu 2011 reported post-operative complication rate; Ellman 2004 reported post-operative complications (other types of complications reported not included in the analysis)



BMJ Open: first published as 10.1136/bmjopen-2018-021967 on 21 September 2018. Downloaded from http://bmjopen.bmj.com/ on June 9, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES)

Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies

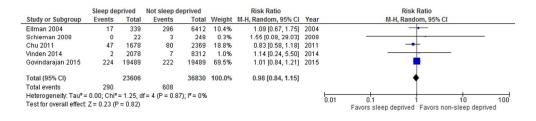


Figure 4. Forest plot for patient mortality among surgeries performed by sleep deprived and non-sleep deprived surgeons

381×101mm (300 × 300 DPI)

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text and

#### **Supplementary file 1.** Search Strategy

Database: In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) 1946 to Present

Date searched: 13 April 2016, updated 7 November 2017

Records retrieved: 5068 and 1442 in the update (removed duplicates retrieved in previous search)

- 1. Medical Staff, Hospital/
- 2. Physician Impairment/
- exp Physicians/
- 4. allergist\*.ti.
- 5. (an?esthetist\* or an?esthesiologist\*).ti.
- 6. cardiologist\*.ti.
- 7. clinician\*.ti.
- 8. (clinician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 9. (clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw,kf.
- 10. dermatologist\*.ti.
- 11. endocrinologist\*.ti.
- 12. doctor\*.ti.
- 13. (doctor\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 14. (doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw,kf.
- 15. family practitioner\*.ti.
- 16. gastroenterologist\*.ti.
- 17. (general practitioner\* or GP\*).ti.
- 18. (general adj2 physician\*).ti.
- 19. geriatrician\*.ti.
- 20. gyn?ecologist\*.ti.

- 22. (health\* adj2 (professional\* or provider\*)).ti.
- 23. hospitalist\*.ti.

- 24. (house staff\* or housestaff\*).ti.
- 25. intensivist\*.ti.
- 26. internist\*.ti.
- 27. medical professional\*.ti.
- 28. obstetrician\*.ti.
- 29. oncologist\*.ti.
- 30. ophthalmologist\*.ti.
- 31. orthop?edist\*.ti.
- 32. (otolaryngologist\* or otorhinolaryngologist\*).ti.
- 33. neonatologist\*.ti.
- 34. nephrologist\*.ti.
- 35. neurologist\*.ti.
- 36. neuropsychiatrist\*.ti.
- 37. neurosurgeon\*.ti.
- 38. p?ediatrician\*.ti.
- 39. perinatologist\*.ti.
- 40. physician\*.ti.
- 41. (physician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 42. (physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw,kf.
- 43. primary care practitioner\*.ti.
- 44. psychiatrist\*.ti.
- 45. pulmonologist\*.ti.
- 46. rheumatologist\*.ti.
- 47. surgeon\*.ti.

- 48. (surgeon\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw,kf.
- 49. traumatologist\*.ti.
- 50. urologist\*.ti.
- 51. or/1-50 [Combined MeSH, title, and text word searches for physicians]
- 52. Burnout, Professional/
- 53. exp Circadian Rhythm/
- 54. exp Fatigue/
- 55. Occupational Health/
- 56. Rest/ph, px [Physiology, Psychology]
- 57. Sleep Deprivation/
- 58. Sleep Disorders, Circadian Rhythm/
- 59. Sleep Wake Disorders/
- 60. exp Stress, Psychological/
- 61. Workload/px [Psychology]
- 62. Work Schedule Tolerance/
- 63. ((24 hour\* or 24 hr\* or twenty four hour\* or twentyfour hour\*) adj rhythm\*).tw,kf.
- 64. biological rhythm\*.tw,kf.
- 65. (burn out\* or burned out\* or burnt out\* or burnout\*).tw,kf.
- 66. circadian misalignment.tw,kf.
- 67. ((circadian or diurnam or ultradian) adj rhythm\*).tw,kf.
- 68. exhaust\*.tw,kf.
- 69. fatigu\*.tw,kf.
- 70. (sleep\* adj3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)).tw,kf.
- 71. tired\*.tw,kf.
- 72. weariness.tw,kf.
- 73. or/52-72 [Combined MeSH and text words for fatigue]
- 74. and/51,73 [Combined concepts for physicians and fatigue]
- 75. animals/ not (animals/ and humans/)
- 76. 74 not 75
- 77. (comment or editorial or letter).pt.

79. limit 78 to yr="2000-Current"

80. limit 79 to (english or french)

81. remove duplicates from 80

Database: Ovid Embase 1996 to 2016 Week 15

Date searched: 13 April 2016

Records retrieved: 8859

- 1. medical staff/
- 2. exp physician/
- 3. allergist\*.ti.
- 4. (an?esthetist\* or an?esthesiologist\*).ti.
- 5. cardiologist\*.ti.
- 6. clinician\*.ti.
- 7. (clinician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 8. (clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 9. dermatologist\*.ti.
- 10. endocrinologist\*.ti.
- 11. doctor\*.ti.
- 12. (doctor\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 13. (doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 14. family practitioner\*.ti.
- 15. gastroenterologist\*.ti.
- 16. (general practitioner\* or GP\*).ti.
- 17. (general adj2 physician\*).ti.

```
18. geriatrician*.ti.
```

- 19. gyn?ecologist\*.ti.
- 20. h?ematologist\*.ti.
- 21. (health\* adj2 (professional\* or provider\*)).ti.
- 22. hospitalist\*.ti.
- 23. (house staff\* or housestaff\*).ti.
- 24. intensivist\*.ti.
- 25. internist\*.ti.
- 26. medical professional\*.ti.
- 27. obstetrician\*.ti.
- 28. oncologist\*.ti.
- 29. ophthalmologist\*.ti.
- 30. orthop?edist\*.ti.
- 31. (otolaryngologist\* or otorhinolaryngologist\*).ti.
- 32. neonatologist\*.ti.
- 33. nephrologist\*.ti.
- 34. neurologist\*.ti.
- 35. neuropsychiatrist\*.ti.
- 36. neurosurgeon\*.ti.
- 37. p?ediatrician\*.ti.
- 38. perinatologist\*.ti.
- 39. physician\*.ti.
- 40. (physician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 41. (physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 42. primary care practitioner\*.ti.
- 43. psychiatrist\*.ti.
- 44. pulmonologist\*.ti.
- 45. rheumatologist\*.ti.
- 46. surgeon\*.ti.

mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

- 48. traumatologist\*.ti.
- 49. urologist\*.ti.
- 50. or/1-49 [Combined Emtree, title, and text word searches for physicians]
- 51. burnout/

- 52. circadian rhythm/
- 53. circadian rhythm sleep disorder/
- 54. fatigue/
- 55. mental stress/
- 56. occupational health/
- 57. sleep deprivation/
- 58. sleep waking cycle/
- 59. work capacity/
- 60. work schedule/
- 61. working time/
- 62. workload/
- 63. ((24 hour\* or 24 hr\* or twenty four hour\* or twentyfour hour\*) adj rhythm\*).tw.
- 64. biological rhythm\*.tw.
- 65. (burn out\* or burned out\* or burnt out\* or burnout\*).tw.
- 66. circadian misalignment.tw.
- 67. ((circadian or diurnam or ultradian) adj rhythm\*).tw.
- 68. exhaust\*.tw.
- 69. fatigu\*.tw.
- 70. (sleep\* adj3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)).tw.
- 71. tired\*.tw.
- 72. weariness.tw.
- 73. or/51-72 [Combined Emtree and text words for fatigue]
- 74. and/50,73 [Combined concepts for physicians and fatigue]
- 75. animals/ not (animals/ and humans/)
- 76. 74 not 75

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

- 77. (conference\* or editorial or letter or proceeding).pt.
- 78. 76 not 77
- 79. limit 78 to yr="2000-Current"
- 80. limit 79 to (english or french)
- 81. limit 80 to embase

Database: Ovid PsycINFO 1987 to April Week 1 2016

Date searched: 13 April 2016

Records retrieved: 2094

- 1. exp Physicians/
- 2. allergist\*.ti.
- 3. (an?esthetist\* or an?esthesiologist\*).ti.
- 4. cardiologist\*.ti.
- 5. clinician\*.ti.
- 6. (clinician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 7. (clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 8. dermatologist\*.ti.
- 9. endocrinologist\*.ti.
- 10. doctor\*.ti.
- 11. (doctor\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 12. (doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 13. family practitioner\*.ti.
- 14. gastroenterologist\*.ti.
- 15. (general practitioner\* or GP\*).ti.
- 16. (general adj2 physician\*).ti.

- 20. (health\* adj2 (professional\* or provider\*)).ti.
- 21. hospitalist\*.ti.

19. h?ematologist\*.ti.

- 22. intensivist\*.ti.
- 23. internist\*.ti.
- 24. medical professional\*.ti.
- 25. obstetrician\*.ti.
- 26. oncologist\*.ti.
- 27. ophthalmologist\*.ti.
- 28. orthop?edist\*.ti.
- 29. (otolaryngologist\* or otorhinolaryngologist\*).ti.
- 30. neonatologist\*.ti.
- 31. nephrologist\*.ti.
- 32. neurologist\*.ti.
- 33. neuropsychiatrist\*.ti.
- 34. neurosurgeon\*.ti.
- 35. p?ediatrician\*.ti.
- 36. perinatologist\*.ti.
- 37. physician\*.ti.
- 38. (physician\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 39. (physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)).tw.
- 40. primary care practitioner\*.ti.
- 41. psychiatrist\*.ti.
- 42. pulmonologist\*.ti.
- 43. rheumatologist\*.ti.
- 44. surgeon\*.ti.

- 45. (surgeon\* adj2 (absent\* or burn out\* or burnout\* or coping or distress\* or duty hour\* or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or well being\* or wellbeing\* or wellness\* or work\* hour\* or work life balance)).tw.
- 46. traumatologist\*.ti.
- 47. urologist\*.ti.
- 48. or/1-47 [Combined thesaurus, title, and text word searches for physicians]
- 49. Compassion Fatigue/
- 50. Fatigue/
- 51. Human Biological Rhythms/
- 52. Occupational Health/
- 53. Occupational Stress/
- 54. Sleep/
- 55. Sleepiness/
- 56. Working Conditions/
- 57. Work Rest Cycles/
- 58. Work Week Length/
- 59. Work Scheduling/
- 60. Workday Shifts/
- 61. ((24 hour\* or 24 hr\* or twenty four hour\* or twentyfour hour\*) adj rhythm\*).tw.
- 62. biological rhythm\*.tw.
- 63. (burn out\* or burned out\* or burnt out\* or burnout\*).tw.
- 64. circadian misalignment.tw.
- 65. ((circadian or diurnam or ultradian) adj rhythm\*).tw.
- 66. exhaust\*.tw.
- 67. fatigu\*.tw.
- 68. (sleep\* adj3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)).tw.
- 69. tired\*.tw.
- 70. weariness.tw.
- 71. or/49-70 [Combined thesaurus and text words for fatigue]
- 72. and/48,71 [Combined concepts for physicians and fatigue]
- 73. limit 72 to yr="2000-Current"
- 74. limit 73 to (english or french)

Date searched: 14 April 2016

Records retrieved: 3378

- S1. (MH "Medical Staff, Hospital+")
- S2. (MH "Physicians+")
- S3. TI allertist\*

- S4. TI (anesthetist\* or anaesthetist\* or anesthesiologist\* or anaesthesiologist\*)
- S5. TI cardiologist\*
- S6. TI clinician\*
- S7. clinician\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or "work\* hour\*" or "work life balance")
- S8. clinician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)
- S9. TI dermatologist\*
- S10. TI endocrinologist\*
- S11. TI doctor\*
- S12. doctor\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or work\* hour\* or "work life balance")
- S13. doctor\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)
- S14. TI "family practitioner\*"
- S15. TI gastroenterologist\*
- S16. TI ("general practitioner\*" or GP\*)
- S17. TI (general N2 physician\*)
- S18. TI geriatrician\*
- S19. TI (gynaecologist\* or gynecologist\*)
- S20. TI (haematologist\* or hematologist\*)
- S21. TI hospitalist\*
- S22. TI ("house staff\*" or housestaff\*)

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

- S23. TI intensivist\*
- S24. TI internist\*
- S25. TI obstetrician\*
- S26. TI oncologist\*
- S27. TI ophthalmologist\*
- S28. TI (orthopaedist\* or orthopedist\*)
- S29. TI (otolaryngologist\* or otorhinolaryngologist\*)
- S30. TI neonatologist\*
- S31. TI nephrologist\*
- S32. TI neurologist\*
- S33. TI neuropsychiatrist\*
- S34. TI neurosurgeon\*
- S35. TI (paediatrician\* OR pediatrician\*)
- S36. TI perinatologist\*
- S37. TI physician\*
- S38. physician\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or "work\* hour\*" or "work life balance")
- S39. physician\* and (cities or city or communit\* or country\* or frontier\* or north\* or remote or rural\* or suburb\* or urban\*)
- S40. TI "primary care practitioner\*"
- S41. TI psychiatrist\*
- S42. TI pulmonologist\*
- S43. TI rheumatologist\*
- S44. TI surgeon\*
- S45. surgeon\* N2 (absent\* or "burn out\*" or burnout\* or coping or distress\* or "duty hour\*" or fatigue or health\* or impair\* or resilien\* or satisfaction or sleep\* or "well being\*" or wellbeing\* or wellness\* or work\* hour\* or "work life balance")
- S46. TI traumatologist\*
- S47. TI urologist\*
- S48. S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28

data mining, Al training, and similar technologies

Protected by copyright, including for uses related to text

S42 OR S43 OR S44 OR S45 OR S46 OR S47

- S49. (MH "Circadian Rhythm")
- S50. (MH "Fatigue")

- S51. (MH "Impairment, Health Professional")
- S52. (MH "Mental Fatigue")
- S53. (MH "Occupational Health")
- S54. (MH "Shiftwork")
- S55. (MH "Sleep Deprivation")
- S56. (MH "Sleep Disorders, Circadian Rhythm")
- S57. (MH "Sleep-Wake Transition Disorders")
- S58. (MH "Stress, Occupational+")
- S59. (MH "Stress, Psychological")
- S60. ("24 hour\*" or "24 hr\*" or "twenty four hour\*" or "twentyfour hour\*") N1 rhythm\*
- S61. "biological rhythm\*"
- S62. "burn out\*" or "burned out\*" or "burnt out\*" or burnout\*
- S63. "circadian misalignment"
- S64. (circadian or diurnam or ultradian) N1 rhythm\*
- S65. exhaust\*
- S66. fatigu\*
- S67. sleep\* N3 (depriv\* or disorder\* or disrupt\* or lack\* or loss or insufficien\* or problem\*)
- S68. tired\*
- S69. weariness
- S70. S49 OR S50 OR S51 OR S52 OR S53 OR S54 OR S55 OR S56 OR S57 OR S58 OR S59 OR S60 OR S61 OR
- S62 OR S63 OR S64 OR S65 OR S66 OR S67 OR S68 OR S69
- S71. S48 AND S70
- S72. S48 AND S70 Limiters Published Date: 20000101-20161231; Publication Type: Clinical Trial, Journal

Article, Meta Analysis, Meta Synthesis, Practice Guidelines, Randomized Controlled Trial, Research,

Review, Systematic Review; Language: English, French

Database: PubMed via NCBI Entrez

Date searched: 14 April 2016

**Records retrieved: 92** 

(((("Medical Staff, Hospital"[mh:noexp] OR "Physician Impairment"[mh:noexp] OR "Physicians"[mh] OR allergist[ti] OR allergists[ti] OR anaesthetist[ti] OR anaesthetists[ti] OR anaesthesiologist[ti] OR anaesthesiologists[ti] OR anesthetist[ti] OR anesthetists[ti] OR anesthesiologist[ti] OR anesthesiologists[ti] OR cardiologist[ti] OR cardiologists[ti] OR clinician[ti] OR clinicians[ti] OR ((clinician[tiab] OR clinicians[tiab]) AND (absent[tiab] OR absentee[tiab] OR absenteeism[tiab] OR absenteeisms[tiab] OR "burned out"[tiab] OR "burn out"[tiab] OR "burn outs"[tiab] OR burnout[tiab] OR burnouts[tiab] OR "burnt out"[tiab] OR coping[tiab] OR distress[tiab] OR distressed[tiab] OR distresses[tiab] OR distressing[tiab] OR "duty hour"[tiab] OR "duty hours"[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR health[tiab] OR healthy[tiab] OR impair[tiab] OR impaired[tiab] OR impairment[tiab] OR resilience[tiab] OR resiliency[tiab] OR resilient[tiab] OR satisfaction[tiab] OR sleep[tiab] OR sleepiness[tiab] OR sleeps[tiab] OR "well being"[tiab] OR wellbeing[tiab] OR wellness[tiab] OR "work hours"[tiab] OR "working hours"[tiab] OR "work life balance"[tiab])) OR ((clinician[tiab] OR clinicians[tiab]) AND (cities[tiab] OR city[tiab] OR community[tiab] OR communities[tiab] OR country[tiab] OR countryside[tiab] OR frontier[tiab] OR north[tiab] OR northern[tiab] OR remote[tiab] OR rural[tiab] OR suburb[tiab] OR suburbs[tiab] OR suburban[tiab] OR urban[tiab] OR urbanite[tiab])) OR dermatologist[ti] OR dermatologists[ti] OR endocrinologist[ti] OR endocrinologists[ti] OR doctor[ti] OR doctors[ti] OR ((doctor[tiab] OR doctors[tiab]) AND (absent[tiab] OR absentee[tiab] OR absenteeism[tiab] OR absenteeisms[tiab] OR "burned out"[tiab] OR "burn out"[tiab] OR "burn outs"[tiab] OR burnout[tiab] OR burnouts[tiab] OR "burnt out"[tiab] OR coping[tiab] OR distress[tiab] OR distressed[tiab] OR distresses[tiab] OR distressing[tiab] OR "duty hour"[tiab] OR "duty hours"[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR health[tiab] OR healthy[tiab] OR impair[tiab] OR impaired[tiab] OR impairment[tiab] OR resilience[tiab] OR resiliency[tiab] OR resilient[tiab] OR satisfaction[tiab] OR sleep[tiab] OR sleepiness[tiab] OR sleeps[tiab] OR "well being"[tiab] OR wellbeing[tiab] OR wellness[tiab] OR "work hours"[tiab] OR "working hours"[tiab] OR "work life balance"[tiab])) OR ((doctor[tiab] OR doctors[tiab]) AND (cities[tiab] OR city[tiab] OR community[tiab] OR communities[tiab] OR country[tiab] OR countryside[tiab] OR frontier[tiab] OR north[tiab] OR northern[tiab] OR remote[tiab] OR rural[tiab] OR suburb[tiab] OR suburbs[tiab] OR suburban[tiab] OR urban[tiab] OR urbanite[tiab])) OR "family practitioner"[ti] OR

"family practitioners"[ti] OR gastroenterologist[ti] OR gastroenterologists[ti] OR "general practice physician"[ti] OR "general practice physicians"[ti] OR "general practitioner"[ti] OR "general practitioners"[ti] OR geriatrician[ti] OR geriatricians[ti] OR gynaecologist[ti] OR gynaecologists[ti] OR gynecologist[ti] OR gynecologists[ti] OR haematologist[ti] OR haematologists[ti] OR hematologist[ti] OR hematologists[ti] OR "health care professional"[ti] OR "health care professionals"[ti] AND "health care provider"[ti] OR "health care providers" OR "health professional"[ti] OR "health professionals"[ti] OR "health provider"[ti] OR "health providers"[ti] OR "healthcare professional"[ti] OR "healthcare professionals"[ti] OR "healthcare provider"[ti] OR "healthcare providers"[ti] OR hospitalist[ti] OR hospitalists[ti] OR "house staff"[ti] OR "house staffs"[ti] OR housestaff[ti] OR housestaffs[ti] OR intensivist[ti] OR intensivists[ti] OR internist[ti] OR internists[ti] OR "medical professional"[ti] OR "medical professionals"[ti] OR obstetrician[ti] OR obstetricians[ti] OR oncologist[ti] OR oncologists[ti] OR ophthalmologist[ti] OR ophthalmologists[ti] OR orthopaedist[ti] OR orthopaedists[ti] OR orthopaedist[ti] OR orthopedists[ti] OR otolaryngologist[ti] OR otolaryngologists[ti] OR otorhinolaryngologist[ti] OR otorhinolaryngologists[ti] OR neonatologist[ti] OR neonatologists[ti] OR nephrologist[ti] OR nephrologists[ti] OR neurologist[ti] OR neurologists[ti] OR neuropsychiatrist[ti] OR neuropsychiatrists[ti] OR neurosurgeon[ti] OR neurosurgeons[ti] OR paediatrician[ti] OR paediatricians[ti] OR pediatrician[ti] OR pediatricians[ti] OR perinatologist[ti] OR perinatologists[ti] OR physicians[ti] OR physicians[ti] OR ((physician[tiab] OR physicians[tiab]) AND (absent[tiab] OR absentee[tiab] OR absenteeism[tiab] OR absenteeisms[tiab] OR "burned out"[tiab] OR "burn out"[tiab] OR "burn outs"[tiab] OR burnout[tiab] OR burnouts[tiab] OR "burnt out"[tiab] OR coping[tiab] OR distress[tiab] OR distressed[tiab] OR distresses[tiab] OR distressing[tiab] OR "duty hour"[tiab] OR "duty hours"[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR health[tiab] OR healthy[tiab] OR impair[tiab] OR impaired[tiab] OR impairment[tiab] OR resilience[tiab] OR resiliency[tiab] OR resilient[tiab] OR satisfaction[tiab] OR sleep[tiab] OR sleepiness[tiab] OR sleeps[tiab] OR "well being"[tiab] OR wellbeing[tiab] OR wellness[tiab] OR "work hours"[tiab] OR "working hours"[tiab] OR "work life balance"[tiab])) OR ((physician[tiab] OR physicians[tiab]) AND (cities[tiab] OR city[tiab] OR community[tiab] OR communities[tiab] OR country[tiab] OR countryside[tiab] OR frontier[tiab] OR north[tiab] OR northern[tiab] OR remote[tiab] OR rural[tiab] OR suburb[tiab] OR suburbs[tiab] OR suburban[tiab] OR urban[tiab] OR urbanite[tiab])) OR "primary care practitioner"[ti] OR "primary care practitioners"[ti] OR psychiatrist[ti] OR psychiatrists[ti] OR pulmonologist[ti] OR pulmonologists[ti] OR rheumatologist[ti] OR rheumatologists[ti] OR surgeon[ti] OR surgeons[ti] OR ((surgeon[tiab] OR surgeons[tiab]) AND (absent[tiab] OR absentee[tiab] OR absenteeism[tiab] OR absenteeisms[tiab] OR

"burned out"[tiab] OR "burn out"[tiab] OR "burn outs"[tiab] OR burnout[tiab] OR burnouts[tiab] OR "burnt out"[tiab] OR coping[tiab] OR distress[tiab] OR distressed[tiab] OR distresses[tiab] OR distressing[tiab] OR "duty hour"[tiab] OR "duty hours"[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR health[tiab] OR healthv[tiab] OR impair[tiab] OR impaired[tiab] OR impairment[tiab] OR resilience[tiab] OR resiliency[tiab] OR resilient[tiab] OR satisfaction[tiab] OR sleep[tiab] OR sleepiness[tiab] OR sleeps[tiab] OR "well being"[tiab] OR wellbeing[tiab] OR wellness[tiab] OR "work hours"[tiab] OR "working hours"[tiab] OR "work life balance"[tiab])) OR traumatologist[ti] OR traumatologists[ti] OR urologist[ti] OR urologists[ti]) AND ("Burnout, Professional"[mh:noexp] OR "Circadian Rhythm"[mh] OR "Fatigue"[mh] OR "Occupational Health"[mh:noexp] OR "Rest/physiology"[mh:noexp] OR "Rest/psychology"[mh:noexp] OR "Sleep Deprivation"[mh:noexp] OR "Sleep Disorders, Circadian Rhythm"[mh:noexp] OR "Stress, Psychological"[mh] OR "Workload/psychology"[mh] OR "Work Schedule Tolerance"[mh:noexp] OR "24 hour rhythm"[tiab] OR "24 hour rhythms"[tiab] OR "24 hr rhythm"[tiab] OR "24 hr rhythms"[tiab] OR alertness[tiab] OR "biological rhythm"[tiab] OR "biological rhythms"[tiab] OR "burn out"[tiab] OR "burned out"[tiab] OR "burnt out"[tiab] OR burnout[tiab] OR "circadian misalignment"[tiab] OR "circadian rhythm"[tiab] OR "circadian rhythms"[tiab] OR "diurnal rhythm"[tiab] OR "diurnal rhythms"[tiab] OR exhausted[tiab] OR exhaustion[tiab] OR exhausting[tiab] OR exhausts[tiab] OR fatigue[tiab] OR fatigued[tiab] OR fatigues[tiab] OR fatiguing[tiab] OR (("Sleep"[mh:noexp] OR sleep[tiab] OR sleeping[tiab]) AND (deprivation[tiab] OR deprive[tiab] OR deprived[tiab] OR deprives[tiab] OR depriving[tiab] OR disorder[tiab] OR disorders[tiab] OR lack[tiab] OR lacked[tiab] OR lacking[tiab] OR lacks[tiab] OR loss[tiab] AND insufficient[tiab] OR problem[tiab] OR problems[tiab])) OR tired[tiab] OR tiredness[tiab] OR "twenty four hour rhythm"[tiab] OR "twenty four hour rhythms"[tiab] OR weariness[tiab] OR "ultradian rhythm"[tiab] OR "ultradian rhythms"[tiab])) NOT (((Animals[MESH] OR Animal Experimentation[MESH] OR "Models, Animal"[MESH] OR Vertebrates[MESH]) NOT (Humans[MESH] OR Human experimentation[MESH])) OR (((animals[tiab] OR animal model[tiab] OR rat[tiab] OR rats[tiab] OR mouse[tiab] OR mice[tiab] OR rabbit[tiab] OR rabbits[tiab] OR pig[tiab] OR pigs[tiab] OR porcine[tiab] OR swine[tiab] OR dog[tiab] OR dogs[tiab] OR hamster[tiab] OR hamsters[tiab] OR chicken[tiab] OR chickens[tiab] OR sheep[tiab]) AND (publisher[sb] OR inprocess[sb] OR pubmednotmedline[sb])) NOT (human[ti] OR humans[ti] OR people[ti] OR children[ti] OR adults[ti] OR seniors[ti] OR patient[ti] OR patients[ti])))) NOT (editorial[pt] OR comment[pt] OR letter[pt] OR newspaper article[pt])) AND ((publisher[sb] NOT pubstatusnihms NOT pubstatuspmcsd NOT pmcbook) OR (pubstatUSheadofprint))

### **Supplementary table 1.** Descriptive characteristics of the included studies

				BMJ Oʻ	5/bmjopen-2018-021967 ccted by copyright, inducention			
Supplementary	table 1. Descriptive char	acterist	ics of the inc	luded studies			8-021 ight, i	
Study	Physician and patient chara	cteristics			Setting		Interventions or S	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban	exposures ding	
						or	<u> </u>	
						rural	for	
Observational (exp	oosure) studies (n=45)						pte En	
Cohort design							mb sei s r	
Chu, 2011 [32]	Surgeons	6	NR	Range: 32-55y	Tertiary care academic	Urban	Insufficient (1997) due	Length of surgery; patient
Canada	Patients: cardiac surgery	4,047	NR	NR	hospital		to work on 🛱 🚔 🤁 t	postoperative mortality,
	cases						preceding s	complications, length of
							Su Su	stay
Ellman, 2004 [41]	Surgeons	NR	NR	NR	University hospitals	Urban	Insufficient \(\mathbb{Q} = \mathbb{Q} \)	Length of surgery; patient
US	Patients: adult cardiac	6,751	70%	S: 63.4±0.7y	<del>-</del>		to work on the hight	complications, in-hospital
	surgery cases			C: 63.5±0.1y			preceding support	mortality, length of stay,
							a ABE	need for blood products
Govindarajan,	Surgeons	1,448	NR	46.3±8.7	Academic and non-	Mixed	Sleep deprivation	Length of surgery; Patient
2015 [31]	Patients: surgical cases	38,978	NR	56.4±16.6y	academic hospitals		to work on 🗖 e n 🙀 ht	complications, mortality,
Canada							preceding a 🗷 ayt 🙀 e	readmissions, length of
							surgery	stay
Rothschild, 2009	Surgeons	220	Surgeons:	Surgeons:	Tertiary care academic	Urban	Sleep deprive tion due	Patient complications,
[58]	Obstetrician/gynecologists		84%	42.0±7.6y	trauma centre/referral		to work on <b>#</b> e n <b>g</b> ht	preventable
US			OB/GYNs:	OB/GYNs:	centre for high-risk		preceding a aytime	complications
			28%	42.0±9.0y	obstetrics		j.com/ on June 9, 202 d similar technologies	
	Patients: surgical and	Surg.:	Surg:	Surg:			com/ on June 9, similar technolo	
	obstetrics cases	4,471	S: 25%	S: 49.1±16.3y			ar t	
		Obst.:	C: 28%	C: 50.0±16.3y			Jui	
		4,902	Obst.:	Obst.:			ne (	
			S: 0%	S: 32.9±5.2y			9, 20: logie	
			C: 0%	C: 33.5±5.0y			Fatigue due to wark	
Schieman, 2007	Colorectal surgeons	NR	NR	NR	University teaching	NR	<b>=</b>	Length of surgery; patient
[63]					hospitals		Ó	operative complications,
Canada	Patients: undergoing	270	NR	S: 64.5y			surgery	length of stay, mortality,
	anterior resection for rectal			C: 64.4y			ě	cancer recurrence
	cancer						3ib	
							surgery ence Bibliographique de	
							<u></u>	
							hic	<u>,</u>
							lue	-
	For	peer rev	iew only - htt	p://bmjopen.b	mj.com/site/about/gu	idelines	xhtml 👱	

				ВМЈ О	oen		6/bmjopen-2018-0	
Study	Physician and patient chara	cteristics			Setting		Interventions or 21	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	including exposures	
Vinden, 2014 [77] Canada	General surgeons Patients: Elective cholecystectomies	331 10,390	83% S: 27% C: 26%	48±10y S: 49±16y C: 49±16y	Community hospitals	Mixed	Sleep deprivation due to overnight with the preceding day with the surgery	Patient mortality, operative complications
<b>Before-after desigr</b> Amirian, 2014 [34] Denmark		29	55%	Median: 35y Range: 27-49y	Academic hospital	Urban	17-h night shift with sleep deprivations	Cognitive and psychomotor abilities or a laparoscopic simulatio
Gerdes, 2008 [45] US	Surgeons	9	NR	NR	University Hospital	Urban	Fatigue; sleep in oad deprivation of the carrier of	Cognitive and psychomotor abilities
Lederer, 2006 [50] Austria	Senior anesthetists	11	82%	49.0±2.0y	Hospital	Urban	Sleep deprive to be from 24-h call she be left.	Concentration ability; reaction time; performance on psychometric tasks
Time series design							rain Bi	
Leichtfried, 2011 [51] Austria	Anesthetists	10	100%	Mean: 32y Range: 29-35y	University Hospital	Urban	Sleep deprivations from 24-h shift; sleepiness, sleepiness, sleepiness sleepi	Melatonin metabolite profile
Cross-sectional des	sign						nilar o	
Aziz, 2004 [35] US	Family medicine physicians Various specialties	153	NR	NR	Hospitals	NR	Fatigue # 3	Stress
Beaujouan, 2005 [36] France	Anesthesiologists	3,476	64%	≤35y: 9% 36-45y: 28% 46-55y: 49% 56-65y: 13%	Public sector General hospitals University hospitals Private hospitals	NR	June echnology, 2025 at Age	Substance abuse
Chang, 2013 [37] US	Anesthesiologists	11	64%	Mean: 38y IQR: 34-48y	Level 1 trauma centre	NR	Sleep deprivation due	Cognitive performance; reaction time
	For	peer rev	riew only - htt	p://bmjopen.b	mj.com/site/about/gu	idelines	shift; sleepiness bij ographique de xhtml	

Study	Physician and patient chara	cteristics			Setting		octed by copyrightein	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures exposures on 2	
Chen, 2008 [38] US	Psychiatrists Internists General practitioners Surgeons Obstetrician-gynecologists Radiologists Pediatricians Other	180	77%	Academic: 79% 36-55y Private practice: 73% 36-65y	Medical school Private practices	Urban	Sleep deprivation Enseignement Supplember 2018. Down sleepiness related to text a	Impact on personal and professional life; perceived risk of errors
Doppia, 2011 [39] France	Anesthesiologists	565	64%	<35y: 11% 35-54y: 63% >55y: 25%	Public hospitals Private hospitals Work-health environments Public health units	NR	sleep deprived data minin	Burnout
Elovaino, 2015 [42] Finland	Physicians in various specialties	1,524	40%	Median: 49.7y Range: 24-69y	Hospitals Primary care Private practice Other unspecified	NR	Sleep difficulties://bmjope	Job demands and contro
Gander, 2000 [43] New Zealand	Anesthetists	183	NR	Mean: 46y	Combined public/private practice Other unspecified	NR	Work-related sleep disturbanced sin	Risk of fatigue-related errors
Harbeck, 2015 [46] Germany	Internists	20	45%	Median: 32y Range: 26-42y	Hospital	NR	Sleep disturbing due to a 24-call wift Jun	Biochemical and physiological parameter neurocognitive function
Heponiemi, 2014 47] Finland	Physicians in various specialties Non-specialized physicians	1,541	40%	49.80±9.49y, Range: 24-67y	Hospitals Primary care clinic Private practice Other unspecified	NR	Sleep difficulties 9, 2025 at Agence Bibliographique de l	Job satisfaction; work ability; psychological distress

		3/bmjopen-2018-0219 cted by copyrightgin							
Study	Physician and patient chara	cteristics			Setting	Intervention	18-021	Outcomes	
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures	67 on 2	
Jackson, 2017 [48] US	Surgeons in various subspecialties	993	61%	More; less satisfied: 30-39y: 23%;24% 40-49y: 32%;36% 50-59y: 23%;27% ≥60y: 23%;14%	Academic practice Non-academic practice	NR		. Download nt Superieu o text and c	Job satisfaction
Kanieta, 2011 [49] Japan	Internists Surgeons Orthopedics Pediatricians Obstetrician-gynecologists Psychiatrists Dermatologists Urologists Opthalmologists Otorhinolaryngologists Other	3,486	66%	20-39y: 11% 40-49y: 25% 50-59y: 28% 60-69y: 16% ≥70y: 21%	Hospitals Clinics Other unspecified	NR	Sleep depri difficulties;	त्व प्रवेत्वात http://bmjopen.bmj.com/ (ABES) रक्षेत्र mining, Al training, and simila	Medical incidents
Lindfors, 2006 [52] Finland	Anesthetists	328	53%	47±7.8y Range: 32-69y	University hospitals Central and district hospitals Private sector	NR		nachnolog	Stress; suicidal tendencies
Mahmood, 2016 [53] Norway	Generalists Internists Pediatricians Surgical specialties Anesthesiologists	450 (all time points)	41%	43y±2.8y	Public health system Private practice	NR	Sleep depri to on-call s	<b>k</b> tion ue	Alcohol misuse
	For	peer rev	iew only - htt	p://bmjopen.b	omj.com/site/about/gu	iidelines	.xhtml	raphique de l	

			5/bmjopen-2018-0219 cted by copyrightsin					
Study	Physician and patient chara	acteristics			Setting		Interventions or 22	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	967 on 2. ncluding	
	Neurosurgeons and neurologists	2,564	NR	NR	Stroke care centres Teaching hospitals	NR	Sleep deprior Enses	Burnout
Pit, 2014 [55] Australia	General practitioners	92	60%	50±10.7y	NR	Rural	Work-related de porton disturbance of contract of the contract	Early retirement intentions
Pit, 2016 [56] Australia	General practitioners	92	60%	50±10.7y	Private (solo) practice Group practice	Rural	Work-related (1999)	Sickness presenteeism
JS	General internists Internal medicine hospitalists	578	58%	Hospitalists: 46.9±12.4y Generalists: 53.6±10.2y	Private practice Academic medical centre Veterans hospital Military practice Other	NR	wnloaded from ht uperieur (ABES) . xt and data mining	Falling asleep while driving
Gaadat, 2016 [60]	Anesthesiologists	21	71%	30-40y: 57% 41-50y: 19% 51-55y: 24% Range: 32-56y	Tertiary care academic children's hospital	NR	Sleep deprivation due to 17-h night call shift aining en. a	
Saadat, 2017 [59]	Anesthesiologists	21	65%	Range: 32-56 years	Tertiary care academic children's hospital	NR	Sleep deprivation due to 17-h nigh call hift	
· · · · · · · · · · · · · · · · · · ·	Emergency medicine physicians	18	28%	29.2±2.6y	Central hospital	NR	Sleep deprivation	Cognitive and psychomotor abilities
Sargent, 2009 [62] JS	Orthopedic surgeons	264	92%	NR	Orthopedic surgery training programs	NR	Sleep deprivations 9, 20	Burnout; psychological distress; marital satisfaction
ende, 2012 [64] France	Emergency physicians	318	62%	39±8y	Hospitals Mobile emergency services Other unspecified	NR	Fatigue; sleep 25 at Agence Bibliographique de xhtml	Stress

		Intervention	i/bmjopen-201 cted by copyr						
Study	Physician and patient cl	haracteristics		Setting			Intervention	ia   8-021	Outcomes
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures	67 on 2	
Sexton, 2001 [65] US	Consulting physicians: Surgeons Anesthesiologists Pulmonary physicians Cardiologists Pediatricians	271	NR	NR	Teaching and non- teaching hospitals	Urban	Fatigue	September 2018. Enseignement	Perceived performance effectiveness
Shanafelt, 2005 [67] US, Canada, Mexico	Oncologists	241	85%	>50y: 51%	Community clinics Hospitals Private practice Academic medical centres	NR	Fatigue; slee		Quality of life/well-being
Shanafelt, 2010 [66] US	Surgeons	7,905	87%	Median: 51y Q1: 43y Q2: 59y	Private practice Academic medical centres Veterans hospital Active military practice Retired or not in practice Other	NR	Fatigue (	om http://bmjopen.bmj.com/ on BES) . mining. Al training, and similar t	Perceived major medical errors
Shanafelt, 2014 [68] US	Oncologists	1,117	52%	Median: 52y	Private practice Academic practice Veteran's hospital Industry, other	NR	Fatigue		Satisfaction with work-life balance
Shirom, 2006 [69] Israel	Opthalmologists Dermatologists Otolaryngologists Gynecologists General surgeons Cardiologists	890	80%	Median: 52y SD: 7.2y	Community clinics Acute care hospital outpatient clinics	NR	Physical fati	June 9, 2025 at Agence Bibliographique de l Sanologies	Perception of quality of patient care
		For peer rev	iew only - htt <sub>l</sub>	o://bmjopen.b	omj.com/site/about/gu	idelines	.xhtml	ographique de l	

			cted by copyrights or	Vbm ionen-2018					
Study	Physician and patient char	racteristics			Setting		Interventions or	Outcome	es
Country	Туре	n=	Sex (% male)	Age	Location	Urban or rural	exposures C C C C C C C C C C C C C C C C C C C	1967 on 2	
Shirom, 2010 [70] Israel	Opthalmologists Dermatologists Otolaryngologists Gynecologists General surgeons Cardiologists	890	80%	Median: 52y SD: 7.2y	Community clinics Acute care hospital outpatient clinics	NR	Physical fatigue Enseignemer de de de la companya d	Burnout	
Smith, 2017 [71] UK	General practitioners Surgeons Other unspecified specialties	3,550	63%	NR	NR (varied)	NR	Perceived fate of sleep deprivation of the sle	Physical health; c	and mental ompetence
Starmer, 2016 [72] US	General pediatricians Pediatric surgeons Pediatric hospitalists Pediatric specialists (unspecified)	840	40%	NR	NR (some in private practice)	NR	Sleep deprivation mining, Al training, and Fatigue	Burnout; personal commitr career sa	balanced and professional nents; life and tisfaction
Tanti, 2017 [73] Malta	Physicians (unspecified)	204	62%	Median: 41y	Hospitals Community Office-based	NR	Fatigue aining	Prescribi	ng errors
Tokuda, 2009 [74] Japan	Hospital physicians: Generalists Other unspecified specialties	236	75%	40.9±7.8y Range: 26-76y	Hospitals with ≥20 inpatient beds	NR O	Sleep deprivation similar to	Burnout;	job satisfaction
Vela-Bueno, 2008 [76] Spain	Primary care physicians	113	27%	41.4±8.0y	Primary care centres	Urban	Sleep problems, insomnia	Burnout	
Wada, 2010 [78] Japan	Physicians (unspecified)	3,862	78%	M: 75% 30- 59y F: 85% 30-59y	Hospitals	NR	Sleep deprivation	Depressi	ve symptoms
							9	raphique	7

				6/bmjopen-2018 cted by copyrig				
Chr. dr.	Dhysisian and nations sh				Cattina		Interventions or 2	Outcomes
Study Country	Physician and patient characteristics Type		Sex (% male)	Λσο.	Setting Location	Urban	- 5 0	Outcomes
Country	туре	n=	Sex (% illale)	Age	Location	or		
						rural	on 2	
Non-comparative o	design						<u> </u>	
Gander, 2008 [43]		20	85%	Median: 44y	Hospitals	Urban	Sleep distur 👼 🙀	Psychomotor
New Zealand				,			from consecutive	performance
							working day	perrermanee
							call work to 20	
ntervention studie	es (n=2)						+ B B	
Randomized contro							o te	
	Emergency physicians	17	35%	39.1y±6.9y	University hospital	Urban		Perceived stress; urine
France	zmergener pmysicians			33.2,20.3,	Cinteroley mospitus	0.20	and 24-h sh	interleukine-8
Tunec							deprivation & deep	
							quality;	
Jchal, 2005 [75]	Surgeons	64	67%	Median:	Government hospitals	NR	Sleep deprive	Product quality,
Norway	Gynecologists	04	0770	Post-call:	Government nospitais	IVIX	to 24-h call ghift	procedure effectiveness
voiway	Orthopedic surgeons			33.0y			to 24-11 can din t	of a surgical simulation
	·						<u>≥</u>	or a surgical simulation
	Urologists Vascular surgeons			Post-work: 38.0y			bmjo train	
Inited Kingdom; US	S: United States of America	; y: year(s)					.bmj.com/ on June 9, 2025 at Agen and similar technologies.	
	F	- For peer rev	iew only - htt <sub>l</sub>	p://bmjopen.k	omj.com/site/about/gu	idelines	2025 at Agence Bibliographique de l gies.	

### **Supplementary file 3.** Risk of bias assessments

Summary of risk of bias assessments for randomized controlled trials (n=2)<sup>a</sup>

First Author, Year	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other sources of bias	Overall risk of bias <sup>b</sup>
Dutheil, 2013	Low	Unclear	High	High	Low	Low	High	High
Uchal, 2005	Low	Low	Unclear	Low	Low	Low	Low	Unclear

<sup>&</sup>lt;sup>a</sup>Assessed using the Cochrane Collaboration's Risk of Bias Tool

Summary of quality assessments for cohort studies (n=6)<sup>a</sup>

First Author,	Selection					Comparab	oility		Outco	me		Total
Year	Representa-	Selection	Ascertain-	Outcome	Total	Compara	Total	Assess-	Adequate	Adequate	Total	Scoreb
	tiveness of	of non-	ment of	not	/4	-bility of	/2	ment of	length of	follow-up	/1	/9
	exposed	exposed	exposure	present at		cohorts		outcome	follow-up	of cohorts		
	cohort	cohort	/1	start		/2		/1	/1	/1		
	/1	/1		/1								
Chu, 2011	1	1	0	1	3	2	2	1)	1	1	3	8
Ellman, 2004	1	1	1	1	4	1	1	1	1	1	3	8
Govindarajan,	1	1	1	1	4	2	2	1	1	1	3	9
2015												
Rothschild, 2009	1	1	1	1	4	2	2	1	1	1	3	9
Schieman, 2008	1	1	1	1	4	1	1	1	1	1	3	8
Vinden, 2014	1	1	1	1	4	1	1	1	1	1	3	8

<sup>&</sup>lt;sup>a</sup>Assessed using the Newcastle-Ottawa Quality Assessment Scale

<sup>&</sup>lt;sup>b</sup>Overall risk of bias is Low if all domains are rated as low, High if at least one domain is assessed as high, and Unclear if at least one domain is assessed as unclear and no domains are assessed as high

<sup>&</sup>lt;sup>b</sup>An overall score of 7 to 9 stars is considered as low risk of bias, 4 to 6 as unclear risk of bias, and 3 or less as high risk of bias

Summary of risk of bias assessments for before-after studies (n=3)<sup>a</sup>

First Author, Year	Random sequence generation <sup>b</sup>	Allocation concealment <sup>b</sup>	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other sources of bias <sup>c</sup>	Overall risk of bias <sup>d</sup>
Amirian, 2014	NA	NA	High	High	Low	Low	High	High
Gerdes, 2008	NA	NA	High	High	Low	Low	High	High
Lederer, 2006	NA	NA	High	High	Low	Low	High	High

<sup>&</sup>lt;sup>a</sup>Assessed using Cochrane Effective Practice and Organization of Care (EPOC) Review Group's criteria for before-after studies, adapted from the Cochrane Collaboration Risk of Bias Tool

## Summary of risk of bias assessments for time series studies (n=1)<sup>a</sup>

First Author, Year	Intervention independent of other changes	Intervention effect pre- specified	Intervention unlikely to affect data collection	Allocation concealment <sup>a</sup>	Incomplete outcome data	Selective reporting	Other sources of bias <sup>c</sup>	Overall risk of bias <sup>d</sup>
Leitchfried, 2011	Low	High	Low	NA	Low	Low	High	High

<sup>&</sup>lt;sup>a</sup>Assessed using Cochrane Effective Practice and Organization of Care (EPOC) Review Group's criteria for interrupted time series studies, adapted from the Cochrane Collaboration Risk of Bias Tool

<sup>&</sup>lt;sup>b</sup>Assessed as 'not applicable' (NA) when the studies did not include a control group

<sup>&</sup>lt;sup>c</sup>Assessed as High due to lack of a control group

<sup>&</sup>lt;sup>d</sup>Overall risk of bias is Low if all domains are rated as low, High if at least one domain is assessed as high, and Unclear if at least one domain is assessed as unclear and no domains are assessed as high

<sup>&</sup>lt;sup>b</sup>Assessed as not applicable (NA) when the studies did not include a control group

<sup>&</sup>lt;sup>c</sup>Assessed as High due to lack of a control group

<sup>&</sup>lt;sup>d</sup>Overall risk of bias is Low if all domains are rated as low, High if at least one domain is assessed as high, and Unclear if at least one domain is assessed as unclear and no domains are assessed as high

Summary of quality assessments for cross-sectional studies (n=34)<sup>a</sup>

First Author, Year		Selection		Outcome				Total Score <sup>b</sup>	
	Adequacy of	Representative-	Total	Assessment of	Same method of	Response rate	Total	/5	
	case definition	ness of the sample	/2	outcome	ascertainment for	/1	/3		
	/1	/1		/1	entire sample				
					/1				
Aziz, 2004	0	0	0	0	1	0	1	1	
Beaujouan, 2005	1	0	1	0	1	0	1	2	
Chang, 2013	1	0	1	0	1	1	2	3	
Chen, 2008	1	0	1	0	1	0	1	2	
Doppia, 2011	1	1	2	0	1	1	2	4	
Elovaino, 2015	1	1	2	0	1	1	2	4	
Gander, 2000	1	1	2	0	1	1	2	4	
Harbeck, 2015	1	0	1	0	1	1	2	3	
Heponiemi, 2014	1	1	2	0	1	1	2	4	
Jackson, 2017	0	0	0	0	1	0	1	1	
Kanieta, 2011	1	0	1	0	1	1	2	3	
Lindfors, 2006	1	1	2	0		1	2	4	
Mahmood, 2017	1	0	1	0	1 //	0	1	2	
Nishimura, 2014	1	1	2	0	1	0	1	3	
Pit, 2014	1	0	1	0	1	1	2	3	
Pit, 2016	1	0	1	0	1	1	2	3	
Roberts, 2014	1	1	2	0	1	0	1	3	
Saadat, 2016	1	1	2	0	1	1	2	4	
Saadat, 2017	1	1	2	0	1	1	2	4	
Sanches, 2015	1	0	1	0	1	0	1	2	
Sargent, 2009	1	0	1	0	1	0	1	2	

First Author, Year		Selection				Total Score <sup>b</sup>		
	Adequacy of case definition	Representative- ness of the sample	Total /2	Assessment of outcome	Same method of ascertainment for	Response rate /1	Total /3	/5
	/1	/1		/1	entire sample			
					/1			
Sende, 2010	1	0	1	0	1	0	1	2
Sexton, 2001	1	0	1	0	1	0	1	2
Shanafelt, 2005	1	0	1	0	1	1	2	3
Shanafelt, 2010	1	1	2	0	1	0	1	3
Shanafelt, 2014	1	0	1	0	1	1	2	3
Shirom, 2006	1	1	2	0	1	1	2	4
Shirom, 2010	1	1	2	0	1	1	2	4
Smith, 2016	1	0	1	0	1	1	2	3
Starmer, 2016	1	1	2	0	1	1	2	4
Tanti, 2017	1	0	1	0	1	0	1	2
Tokuda, 2009	1	1	2	0	1	1	2	4
Vela-Bueno, 2008	1	1	2	0	1	1	2	4
Wada, 2010	1	1	2	0	1	0	1	3

<sup>&</sup>lt;sup>a</sup>Assessed using the Newcastle-Ottawa Quality Assessment Scale, adapted for cross-sectional studies

<sup>&</sup>lt;sup>b</sup>An overall score of 4 to 5 stars is considered as low risk of bias, 3 as unclear risk of bias, and 2 or less as high risk of bias. For response rate, ≥50% was used as the criterion to be awarded a star

Summary of quality assessments for non-comparative studies (n=1)<sup>a</sup>

First Author,		Selection		Exposu	re		Outcome			Total
Year	Adequacy of case definition /1	Representat- iveness of the sample /1	Total /2	Ascertain- ment of exposure	Total /1	Assessment of outcome /1	Same method of assessment for entire sample	Loss to follow-up /1	Total /3	Score <sup>b</sup> /6
Gander, 2008	1	1	2	0	0	0	<b>/1</b> 1	1	2	4

<sup>&</sup>lt;sup>a</sup>Assessed using the Newcastle-Ottawa Quality Assessment Scale, adapted by the authors to be suitable to the non-comparative design

<sup>&</sup>lt;sup>b</sup>An overall score of 5 to 6 stars is considered as low risk of bias, 3 to 4 as unclear risk of bias, and 2 or less as high risk of bias 

# Physician health and wellness outcomes and associations with fatigue

Study	Study	<b>Exposures or interventions</b>		Outcomes	Associations Between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	ept
		time points		time points	ptem Ens. uses
Surgeons					ber rela
Jackson, 2017	CS	Not feeling well rested: self-	71% healthy, 28% unhealthy in	Job satisfaction: Abridged	Job satisfaction hose more vs. less satisfied:
		reported as 'unhealthy'	terms of being well rested	Job in General Scale;	Healthy (welk 🚉 🗪 ): 85% vs. 58%, p<0001;
RoB: high				grouped into more or less	Unhealthy (ngt 📆 🖁 rested): 15% vs. 42%, p<0.003
		Time points NR		satisfied using the median	wni xt a
				Time points NR	nd da
Nishimura, 2014	CS	Sleep hours/night: self-	Mean±SD sleep: 5.94±1.08h	Burnout: Japanese MBI	1) Mean±SD seep or not burned out vs. mild to
		reported (continuous)		(severe: EE >4.0 and either	moderate vs 3 7 8 6 6.07 ±1.15 vs. 5.88 ±0.94 vs.
RoB: unclear				DP >2.6 or PE <4.17)	5.63±0.94, p <b>3</b> 0.05 <del>2</del>
		Time points NR			2) Association between sleep and burnout (OR
				Time points NR	(95% CI)): bių (0.61-0.73), p<0.001;
					multivariate 🖺 clu <mark>a</mark> ng work characteristics and
					mental healt (0.75-0.94), p=0.002.
Sargent, 2009	CS	Sleep deprivation: self-	21% none, 48% a little, 23%	Burnout: MBI (norms NR);	1) Positive correlation between sleep deprivation
		reported on a 4-point scale	quite a bit, 8% a lot	Marital satisfaction: RDAS;	and EE, DP, psychological distress, lower marital
RoB: high		(none, a little, quite a bit, a		Psychological morbidity:	satisfaction, 省 p💁.001. No relationship with PA.
		lot)		GHQ-12 score ≥4	n/ ou
		Time points NR		Time points NR	n June techno
Anesthesiologists <sup>a</sup>					noi-
Lederer, 2006	BA	24-h shift with on-call duty;	Mean±SD sleep: 4.1±1.7h;	Stress during duty: 4-point	1) Mean stress score during duty: 2.1.
		Sleep hours and	Number of interruptions:	scale from 'calm' to 'very	»s. 25
RoB: high		interruptions: self-reported;	0.8±1.1;	demanding'	at t
		Tiredness: VAS from 0 (low)	Tiredness pre- vs. post-duty:		Age
		to 100 (high)	30.9±27.5 vs. 59.5±18.9,	Assessed post-duty	Agence Bibliogra
			p=0.01.		Ф Ш
		Assessed pre- and post-duty			<u>5</u>

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and time points	Baseline	Assessment measure and time points	967 o
Leitchtfried, 2011	TS	24-h shift;	ESS (mean (range)): 7.4 (4-12);	aMT6-s: urinalysis	1) aMT6-s over shift, mean (95% CI): higher at
		Sleepiness: ESS (range: 0-	Mean±SD sleep hours:		11:00AM pre (12@ (6.3-8.1)) and post-shift (9.3
RoB: high		24);	1) pre-study: 7.74±1.35h;	Assessed at 4-h intervals	(3.7-14.9)) v <b>与</b> c山道g, p=0.016;
		Sleep hours: self-reported	2) Pre-24-h shift (11h00 on day	from 07:00 to 11:00	2) Correlations de ween sleep and aMT6-s (data
		(continuous)	1: 0.13±0.35h, 19:00 on day 1:		NR): mild for ( ) with a sign of the night prior with
			6.99±0.68h);		aMT6-s at 3PM shellowing day; sleep on night 2
		Sleepiness assessed pre-shit,	3) During the 24-h shift (07h00		with aMT6-s and the next day; total sleep with
		sleep hours pre, during and post-shift	on day 2: 0.0±0.0h, 19h00 on day 2, 5.49±1.95h);		aMT6-s at 11% W third day; moderate for sleep
	post-snirt	4) Post-24-h shift (11h00 on		on first night with \$MT6-s at 7AM and 11AM preshift, 11PM @ 2022 24-h shift and 11AM post-shift	
		day 3: 0.5±0.71h, 19h00 on		total sleep pe 与最 and nocturnal sleep during 24	
			day 3: 7.06±1.18h).		h shift with a sat 11PM during shift; total
					sleep with al 2 5 at 3PM on first and second day
					11PM on second day;
					3) Correlations between ESS and aMT6-s:
					moderate for MB-s at 7AM during shift, 11AM
					on day off. 👜 👼
Beaujouan, 2005	CS	Sleep deprivation: 4-point	48.8% always or frequently	Substance abuse: 93-item	1) 60.6% wit dependence vs. 46.0% of those
		scale (always, frequently,	feel sleep deprived	addiction and substance	without reported beep difficulties, p<0.001.
RoB: high		rarely, never)		abuse questionnaire	2) OR (95% Cad of addiction for frequently/always
		Time mainte ND		Time a points ND	vs. rarely/ne or sleep deprived: tobacco 1.42
		Time points NR		Time points NR	(1.04-1.94); tanquilizer/hypnotics 3.26 (2.12-5.02).
Doppia, 2011	CS	Insufficient sleep: 4-point	28.9% reported insufficient	Burnout: CBI (mild: 1-2.4,	1) Frequency of be nout by response for sleep
		scale (no, not really, sort of,	sleep during work time	moderate: 2.5-3.5, severe:	sufficiency: 42.6% or no/not really, 16.3% for sor
RoB: low		yes)		3.6-5)	of/yes, p<0.
		Time points NR		Time points NR	9S.
Lindfors, 2006	CS	Sleep hours/day: self-	Sleep hours (mean (range)): 7	Stress: MOSQ on a 3-point	1) Sleep sufficience predicted stress symptoms:
		reported to the nearest	(5-9)	scale (no, to some extent,	bivariate $\beta$ =-0.36 $\frac{2}{3}$ p<0.001; multivariate including
RoB: low		0.5h;		clearly);	gender, sick leave suicide $\beta$ =-0.269, p<0.001;
		Adequacy of sleep and rest:		Thoughts of suicide: 4-point	2) Sleep disturbange associated with thoughts of
		self-reported (yes/no)		scale ('never' to 'have tried')	suicide, p=0.009. og ap phi que s.xhtml de

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and time points	Baseline	Assessment measure and time points	967 on
		Time points NR		Time points NR	21 S
Saadat, 2015	CS	Sleep deprivation (<7h/24-h)	Mean±SD sleepiness on a	Simple cognitive tests: VAS	Regular day र्रामुक्ट call day, mean±SD scores:
		due to 17-h overnight shift;	regular day vs. post-call day:	from 0 (not at all) to 100	1) Simple coshigwes tests: energetic 6.04±2.27 vs
RoB: low		Sleepiness and alertness:	2.99±2.18 vs. 6.79±2.30,	(extremely); Mood disturbance: PMS	2.53±1.87, com and t 7.03±1.83 vs. 4.98±2.29,
		VAS from 0 (not at all) to 100 (extremely)	p<0.001	(scoring NR)	irritable 2.03 (2015) s. 4.86±2.16, sleepy 2.99±2. vs. 6.79±2.30 (2015) sive 4.46±1.74 vs. 2.41±1.97,
		100 (extremely)			p<0.001; jitter 2.4.4011.74 vs. 2.4111.97,
		All assessed on a regular day		All assessed on a regular day	anxiousness #55 \$
				and a post-call day	2) PMS: tenser 1 & 48±2.71 vs. 15.43±4.46,
				,	p=0.049; anger = 52 4±4.41 vs. 18.14±5.92,
					p=0.005; fati <b>gu p</b> 14±2.63 vs. 20.05±6.87,
					p<0.001; con 10.57±1.69 vs. 12.57±4.24,
					p=0.025; viges 24.25±6.75 vs.16.67±5.70, p<0.00
					depression: ns; total mood disturbance:
				<b>•</b>	42.57±15.26 xs. 7690±6.91, p<0.001.
ER or ICU physician		441 241 1:6:	4) ()	C	1) C) 1: E' (9) : 11   121   15
Dutheil, 2013	RCT	14-h or 24-h shift; Sleep hours: self-reported	1) Sleep duration and quality lower during shifts (14h and	Stress: VAS from 0 (low) to 100 (high);	1) Stress: higger following 14-h and 24-h shifts v the control day, p 0.05 (data NR);
RoB: high		sleep and wake time;	24h) than any other day, and	IL-8: urinalysis	2) IL-8: highed following 24-h shift vs. control
NOD. HIGH		Sleep quality: VAS from 1	lower during the 24-h vs. 14-h	TE O. diffidiyala	(p=0.007) an $\frac{2}{5}$ shift (p=0.015); ns difference
		(low) to 100 (high);	shift (p<0.05);	Assessed at 08:30 and 18:30	between 14- Shift and control day;
		Mental and physical fatigue:	2) Mental and physical fatigue	on each day of protocol	3) Correlations with IL-8: sleep hours pre-24-h
		VAS from 1 (low) to 100	higher after 14-h and 24-h shift		shift, r=-0.62 p= 007; poor sleep quality during
		(high)	vs. control day (data NR).		14-h and 24- shifes, r=0.452, p=0.031;
					4) Multivaria le regression: 24-h shift increased
		Assessed on day prior to			8 by 1.9ng vs. con olday, p=0.007; ns association
		shift; during shift; each day			with 14-h shift, mental or physical fatigue, sleep
		of protocol (work, off,			deprivation, 14-h
Sanda 2012	CC	clerical, control)	NR	Most important	1) 78% indicated that sleep loss and fatigue wer
	CS	Fatigue and sleep	INIC	Most important sources of	sources of stress.
Sende, 2012		deprivation as sources of		stress among 4 categories	COLLECTE OF CTRACE

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	
		time points		time points	ā. o
				related, organizational,	ng :
		Time points NR		individual)	for Se
				Time points NR	n 21 September 201 Enseigneme
Generalists <sup>b</sup>					nr 20 Jiner ate
Harbeck, 2015	CS	24-hours on-call shift with	1) Sleep hours on a normal day	Biochemical (laboratory	Before a norgaes s 解ft vs. after overnight call shift:
		sleep disturbance: self-	vs. following a 24-h shift:	values) and physiological	1) Biochemica pameters: no changes in any
RoB: unclear		reported number of sleep	<2 hours: 0 vs. 5.9%; 2-4	(heart rate variability, skin	parameter eॡ्कें क्वाई_or thyroid stimulating hormon
		disturbances and hours of	hours: 5.9% vs. 47.1%; 4-6	resistance, blood pressure)	which was high after the on-call shift (p = 0.049
		sleep per night	hours: 11.8% vs. 35.3%; >6	stress parameters	data NR); da i de
			hours: 82.4% vs. 11.8%		2) Physiologien Parameters: no significant change
		Assessed before a normal	2) Number of sleep	Assessed before a normal	in any paramaty
		day shift, and after a 24-h on call shift	disturbances a normal day vs.	day shift, and after a 24-h on call shift	ng, · http
		Call Stillt	following a 24-h shift: 0: 82.4% vs. 11.8%; 1: 11.8%	Off Call Stiff	Al 1
			vs. 35.3%; 2: 5.9% vs. 47.1%; 3:		raii mj
			0% vs. 5.9%; 4: 0% vs. 0%; >4:		nttp://bmjopen. ng, Al training,
			0% vs. 0%		g, n. a b
Pit, 2014	CS	Work-related sleep	Work-related sleep	Early retirement (<65 years)	For sleep distance a few times a year to every
		disturbance: 7-point scale	disturbance: 41% never, 59% a	intentions (yes/no)	day vs. never
RoB: unclear		from 'never' to 'every day'	few times a year to every day		1) Intention a retire early: 74% vs. 26%, p<0.01;
		Time and into NID		Time points NR	2) Association with intention to retire early (OR
		Time points NR			(95% CI)): ungariage 3.6 (1.47-8.80), p<0.01; multivariate Octubracy work, occupational,
					individual factors & 91 (1.11-7.6), p<0.05;
					4) RR (95% Cl) for tention to retire early: 2.0
					(1.18-3.49); attribation: 50.0%;
					population attributable fraction: 37.1%.
Pit, 2016	CS	Work-related sleep	Work-related sleep	Sickness presenteeism: 'yes'	For sleep disturbæce a few times a year to every
		disturbance: 7-point scale	disturbance: 41% never, 59% a	response indicated 1 or	day vs. never: 👸
RoB: unclear		from 'never' to 'every day'	few times a year to every day	more days	1) Sickness preserteism: 32% vs. 68%, p=0.018;

cted by copyrigh 6/bmjopen-2018-0

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	967 o
		time points		time points	udi o
		Time points NR		Assessed for the past 12	2) Association with sickness presenteeism (OR
				months	(95% CI)): 2.92 (1.40-7.16), p=0.02.
Roberts, 2014	CS	Fatigue: LAS from 0 (low) to	Mean (SD) score: 5.8 (2.4) for	Impact of fatigue on daily	1) 8.7% of ho
		10 (high)	hospitalists; 5.9 (2.4) for	activities (falling asleep	general inter இத்து ad fallen asleep while driving
RoB: unclear			general internists	while driving) (yes/no)	due to fatigue o
		Assessed for the past week			20 ate
				Time points NR	due to fatigue ated to
Vela-Bueno, 2008	CS	Sleep Quality: PSQI	Prevalence (% (95% CI)):	Burnout: PBM with a 7-point	Low vs. high கும் திர், mean±SD:
		(Spanish): score ≥5 indicates	1) Sleep-onset latency >30	scale from 1 (never) to 7	1) Global PS(한 등 출±2.22 vs. 7.24±4.17, p<0.001;
RoB: low		low quality (range; 0 to 21);	minutes: 8.4 (4.8-11.9);	(always)	2) PSQI subs 🛱 🔓 s sileep quality: 0.54±0.57 vs.
		Insomnia: DSM-IV criteria	2) Wake time after sleep onset		1.40±0.83, p 🕶 🗓 🕏, sleep latency: 0.51±0.80 vs.
			>30 minutes: 15.4 (10.8-19.9);	Time points NR	1.38±1.03, p 2 sleep duration: 0.45±0.64 vs.
		Time points NR; insomnia	3) Early morning awakening:		1.16±0.92, p ; sleep efficiency: 0.21±0.57 vs
		symptoms in past month	22.5 (19.5-30.4);		0.77±0.98, p=0.012; sleep disturbance: ns; use of
			4) Nonrestorative sleep: 22.5		medication: 0.14±0.49 vs. 0.57±0.83, p=0.032;
			(17.2-27.7);		daytime dysfanction: 0.52±0.73 vs. 1.57±0.88,
			5) Daytime impairment for ≥5		p=0.002.
			days in past month: 14.2 (9.7-		3) Prevalenc (95 CI) of insomnia symptoms:
			18.6);		sleep latency 5.5% (2.5-11.5%) vs. 21.1% (10.5-
			6) Insomnia: 18.8 (13.8-23.7).		31.6%), p=0. 3:5; ake time >30 min after sleep
					onset: 9.4% 😩 6-1 <mark>2</mark> .1%) vs. 25.5% (14.2-37.7%),
					p=0.029; ear awakening: 14.5% (5.1-23.8%) vs.
					45.6 (32.7-58,4%) p<0.001; somewhat/very
					dissatisfied with seep: 5.5% (2.5-11.5%) vs. 50%
					(37.1-62.8%) o <0.001; day impairment: 5.5% (2.5-
					11.5%) vs. 3 2 % (25.6-50.7%), p<0.001; insomnia
					7.3% (0.4-14 <b>%</b> ) vs <mark>\$</mark> 39.7% (27.1-52.2%), p<0.001.
Oncologists					<u>. 01</u>
Shanafelt, 2005	CS	Fatigue: LASA QOL ≤7;	75% had a high level of fatigue;	Wellbeing: 10-item LASA	1) Sleep deprivatian for high vs. low overall well-
		Sleep deprivation: 10-point	Mean±SD sleep score:	QOL, high ≥8 vs. low ≤7	being (mean±SD): 3.9±2.57 vs. 5.1±2.60, p=0.0004
RoB: unclear		Likert scale from 0 (not at	4.5±2.65.		2) Lower fatigue pedicted overall wellbeing in a
		all) to 10 (stressful as can		Time points NR	multivariate mod fincluding personal and
		be)			professional characteristics, p=0.002.
					graphi
					que .
		For peer revi	ew only - http://bmjopen.bmj	j.com/site/about/guideline:	graphique de

Study Stud Risk of Bias (RoB) designment Shanafelt, 2014 CS RoB: unclear  Mixed groups of physicia Aziz, 2004 CS	Exposures or interventions Assessment measure and time points  Time points NR Fatigue: 10-point LAS (lower scores indicate greater fatigue)  Time points NR	Mean±SD fatigue score: 5.7±2.4	Assessment measure and time points  Satisfaction with WLB: 5-point Likert scale from 'strongly agree' to 'strongly	Associations between exposure and outcome  97 97 97 97 97 97 97 97 97 97 97 97 97
RoB: unclear  Mixed groups of physicia	Time points NR Fatigue: 10-point LAS (lower scores indicate greater fatigue)		Satisfaction with WLB: 5- point Likert scale from	ပြု ညီ ရှိ ဖွာ 1) OR (95%Cနှင့်ရှိမြွှေwer satisfaction predicted by
RoB: unclear  Wixed groups of physicia	Fatigue: 10-point LAS (lower scores indicate greater fatigue)		point Likert scale from	ပြု ညီ ရှိ တွ 1) OR (95%Cနှင့်ရှိမြွှဲwer satisfaction predicted by
RoB: unclear  Wixed groups of physicia	Fatigue: 10-point LAS (lower scores indicate greater fatigue)		point Likert scale from	1) OR (95%C (Garage of the Section 1) OR (95%C (Garage of the Section 2) OR (95%C (Gar
RoB: unclear  Wixed groups of physicia	scores indicate greater fatigue)		point Likert scale from	
Mixed groups of physicia	fatigue)	5./±2.4	•	high fatigue (ws@ngt) in multivariate model
Mixed groups of physicia			'strongly agree' to 'strongly	
	Time points NR			including personal and work-related factors, and
	Time points NR		disagree'	burnout: 0.4
			Time naints ND	8. C to 1
	ne		Time points NR	
1212, 2001	Working while fatigued: 5-	NR	Stress: 47-item	1) Sources of the second will be second to the second seco
	point scale from 'extreme'		questionnaire with a 5-point	— v =
RoB: high	to 'a little'		scale from 'extreme' to 'a	0.653, in factor analysis;
			little'	2) Inverse co
	Time points NR			while fatigue: 1= 2.270 (significance level NR).
	•		Time points NR	6, 7, 6.
Chen, 2008 CS	Sleepiness: ESS score ≥11	Mean±SD ESS score: 7.8±4.0,	Impact on work and	1) Impact score correlated with ESS, r=0.31,
		range: 0-20, 23% had scores	personal life: Impact	p<0.05; <b>a</b> i <b>j</b> o
RoB: high	Time points NR	≥11.	Questionnaire with a 5-	2) ESS score among physicians who
			point Likert scale from 1	agree/strong agee vs. other response: worried
			(strongly agree) to 5	about having caraccident while driving home
			(strongly disagree)	post-call: $5.4\frac{8}{2}$ s. $7\frac{1}{2}$ , p<0.001; sleep loss has a
				major impaction personal life: 8.4 vs. 7.0, p=0.01
			Time points NR	3) Higher ESS cores predicted by impact score in
				multivariate gression including personal and
				work-related dact control state (β s: β=0.11, p=0.005.
Elovaino, 2015 CS	Sleeping problems: Jenkins	Mean±SD score:	Jobs demands: 5 items	There was neassoniation between sleeping
D-D-I	Scale with a 6-point scale	2006: 2.30 (1.00);	scored on a 5-point scale	problems in 3006 and job demands or control in
RoB: low	from 1 (never) to 6 (every	2010: 2.35 (1.05).	from 1 (strongly disagree) to	2010. at Agence Bibliographique de l
	night)		5 (strongly agree); Job control: 3 items derived	g e
	Assessed in 2006 and 2010		from the Karasek Job	nce
	Assessed III 2000 and 2010		Questionnaire	m

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and time points	Baseline	Assessment measure and time points	967 or
Heponiemi, 2014	CS	Sleeping problems: Jenkins Scale <sup>81</sup> with a 6-point scale	Mean±SD (range) score: 2.30±1.00 (1-6)	Psychological distress: GHQ-	1) Sleeping problems associated with job satisfaction, <b>2</b> =-0. <b>42</b> , p<0.001, psychological
RoB: low		from 1 (never) to 6 (every	2.30±1.00 (1-0)	12 with a 4-point scale (low to high);	satisfaction, $\mathbf{g} = 0.001$ , psychological distress, $\beta = 0.001$ , $\mathbf{g} = 0.001$ ;
1105.1011		night)		Job satisfaction: JDS with a	2) Total indirection because the second control of the second cont
				Likert scale from 1 (strongly	mediators (s
		Assessed in 2006		disagree) to 5 (strongly	with family) 👸 📆 🗫 CI)): job satisfaction 0.06 (-
				agree)	0.059, -0.016, <b>1</b> 0.001; psychological distress 0.
				Assessed in 2010	(0.023, 0.081) 6 0001.
Mahmood, 2016	CS	Sleep deprivation: self-	Mean±SD hours:	Alcohol use disorders:	There was near seem to sleep
		reported mean hours of	4 years: 4.52 (2.79);	Modified 9-item version of	when on call (Range ) Bazardous drinking behaviours
RoB: high		sleep when on call	10 years: 5.38 (6.36);	the Alcohol Use Disorder	(p=0.732) a 2 2 3
-		15 years: 6.41 (7.14).	Identification Test (AUDIT)	m BEC	
		Assessed at 4 years, 10		≥6 for men and ≥5 for	om http 3ES) · mining,
		years, and 15 years post-		women.	, AI .//
		graduation		Assessed at 4 years, 10	http://bmjopen ) . ing, Al training
				years, and 15 years post-	
				graduation	g, a
Shirom, 2010	CS	Tiredness and exhaustion:	NR	Burnout: SMBM on a 7-	1) Correlatio between physical fatigue subscale
		SMBM Physician Fatigue		point scale from 1 (almost	and overall burnogt: 0.88, p<0.05;
RoB: low		Subscale on a 7-point scale		never) to 7 (always)	2) In a predictive structural model for burnout,
		from 1 (almost never) to 7			physical fatigue accounted for unique variance in
		(always)			the burnout \(\mathbb{G}\)ems_not accounted for by total burnout (\(\mathbb{R}^2 = \overline{\overline{B}}\)24)
		Time points NR			0 0 0 2
Smith, 2017	CS	Sleep deprivation: self-	NR	Mental and physical illness:	Some physicans reported developing mental
		reported via open-ended		self-reported via open-	illness (e.g., bipol disorder, alcohol misuse) due
RoB: unclear		comments		ended comments	to tiredness and seess at work; others developed
		Time a resinta ND		Time a ciuta ND	physical health presidents due to sleep deprivation
		Time points NR		Time points NR	poor eating habit and lack of exercise.

Study	Study	Exposures or interventions		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and time points	Baseline	Assessment measure and time points	967 or
Starmer, 2016 RoB: low	CS	Sleep deprivation: <7 hours sleep in a typical 24-h period (self-reported)	27.7% sleep deprived	Burnout, satisfaction with career and life, balanced personal and professional commitments: Each on a 5-point Likert scale (strongly	≥7-h vs. <7-hos leeps  1) Burnout (? stropgly agree/agree): 26.4% vs. 39.6%, p<0.05; mileer satisfaction (% strongly agree/agree) for set is faction (% completely/ 2 set is fied): 76.4% vs. 55.9%, p<0.05; balance agree ≥ 2 set is faction (% completely).
				agree to strongly disagree) Time points NR	commitments of the same personal and professional commitments of the same personal factor and professional and professional commitments of the same personal factor and professional and professional commitments of the same personal factor and professional commitments of the same
Tokuda, 2009 RoB: low	CS	Sleep hours/day: self-reported (continuous)  Time points NR (included weekday and weekends)	Mean±SD (range) sleep hours/day: 6±0.9 (3-8)	Burnout: MBI (Japanese) with a 7-point Likert scale: 0 (none) to 6 (every day); Job satisfaction: JHPSS with a 5-point Likert scale: 1 (strongly disagree) to 5 (strongly agree)	Maximum likelihood estimates±SE:  1) Sleeping time to job satisfaction: group 0.990±0.458 = 0.031; ns for men; women 1.711±0.805 = 0.034; 2) Sleeping time to job satisfaction: group 0.990±0.458 = 0.031; ns for men; women 1.711±0.805 = 0.034; 2) Sleeping time to job satisfaction: group 0.990±0.458 = 0.031; ns for men; women 1.711±0.805 = 0.031; ns for men; women 1.711
Wada, 2010 RoB: unclear	CS	Sleep hours/day: Self-reported (continuous)  Assessed for past month when not completing overnight work	<5 hours: 8.7% men, 9.9% women; 5 to <6 hours: 32.3% men, 34.6% women; 6 to <7 hours: 46.0% men, 43.7% women; ≥7 hours: 13.0% men, 11.8% women.	Time points NR  Depression: QIDS-SR; Japanese score <5 (no symptoms) to >20 (very severe symptoms)  Assessed for past 7 days	1) Sleep house for hose with vs. without depressive sumptions: <5: 18.7% vs. 7.7% men, 20.5% vs. 8.7% we hen; 5 to <6: 33.7% vs. 32.2% men, 38.6% vs. 342% women; 6 to <7: 35.1% vs. 46.9% men; 31.8% vs. 45.1% women; 2) Association between <5h sleep (vs. 6-7h) and depressive symptoms (OR (95% CI)): univariate 2.79 (1.96-3.95) for men, 2.65 (1.47-4.78) for women; multivariate (including age and workload)

					<del></del> <u> </u>
Study	Study	Exposures or interventions		Outcomes	Associations <u>betyleen</u> exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	- 967
		time points		time points	udi
					ft 2 70 74 02 4 02) f 2 20 /4 44 F 40)

factors) 2.7061.824.03) for men, 2.38 (1.11-5.10)

for women. **ਰੁੱ** 

cted by copyrig 6/bmjopen-2018

alncludes studies of anesthetists, where these were physicians.

blincludes primary care physicians, internal medicine physicians, and general practitioners.

AM: morning; aMT6-s: melatonin metabolite; BA: before-after; CI: confidence interval; CBI: Copenhagen Burnout Inventory; CS: cross-sect Diagnostic and Statistical Manual of Mental Disorders; EE: emotional exhaustion; ER: emergency; ESS: Epworth Sleepiness Scale; GHQ: Geriging Realth Questionnaire; h: hour(s); ICU: intensive care unit; IL-8: interleukin-8; JDS: Job Diagnostic Survey; JHPSS: Japanese Hospital Physicians Satisfaction Scale; LAS: linear aparticles and logical control of the contr assessment scales; MBI: Maslach Burnout Inventory; MOSQ: Modified Occupational Stress Questionnaire; min: minute(s); NA: not applicate; not reported; ns: not statistically significant; OR: odds ratio; PA: personal achievement; PBM: Pines Burnout Measure; PE: professional efficacy; PM: afternoon; 👫 Frofile of Mood States; PSQI: statistically significant; OR: odds ratio; PA: personal achievement; PBM: Pines Burnout Measure; PE: professional efficacy; PM: afternoon; PM: grofile of Mood States; PSQ:: Pittsburgh Steep Quality Index (2DS-SR: Quilek Inventory Depressive Scale - Self-Reported; QOL: Quality of Life; RCT: randomized controlling and significant; PM: grofile of Mood States; PSQ:: Pittsburgh Steep Quality Index (2DS-SR: Quilek Viewford) Ppt (2D) and significant of Life; RCT: randomized controlling and significant of Life; Pittsburgh Sleep Quality Index; QIDS-SR: Quick Inventory Depressive Scale – Self-Reported; QOL: Quality of Life; RCT: randomized controlled that RDAS: Revised Dyadic

cted by copyright, 3/bmjopen-2018-02

Performance and safety outcomes related to fatigue or sleep loss among physicians in independent practice

Study	Study	Exposures or intervention		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	on 2
		time points		time points	າg <sub>21</sub>
Surgeons					or se
Uchal, 2005	RCT	Sleep deprivation from a 24-	Median (range) sleep hours:	Surgical performance:	Post call va post-work:
		h call shift vs. 8-h work;	1.5 (0-3) post-call vs. 6.5 (5-	laparoscopic surgical	1) Producto
RoB: unclear		Sleep hours: self-reported	9) post-work, p<0.05;	simulator(Minimally Invasivs	error, tiss 🏖 🗗 🏝 age, leak rate;
		(continuous);	Median ESS score: 7.0 post-	Surgical Trainer-Virtual	2) Proced 🛱 🕰 ectiveness: no difference in
		Sleepiness: ESS (moderate:	call vs. 5.5 post-work, ns.	Reality) for product quality,	goal-direce actions, non-goal directed
		10-15, severe: ≥16)		procedure effectiveness	actions, of ecology time.
		Assessed post-call and post-		Assessed post-call and post-	nloade perieur
		work		work	W — Ö
Chu, 2011	СО	Sleep deprivation: self-	Of 4,047 procedures, 83	Surgical performance: CABG,	For 0-3 vs 3 5 5. >6 hours of sleep: no
		reported hours, moderate	(2.1%) performed by	ACC	difference BBG or ACC.
RoB: low		(3-6h) or severe (<3h)	severely sleep-deprived and		=:-C <u>-</u>
			1,595 (39.4%) moderately	Assessed during surgery	<b>5</b>
		Assessed the night before	sleep-deprived surgeons		://bmjo Al train
		surgery			<b>≡.</b> →
Ellman, 2004	СО	Sleep deprivation:	Of 6,751 procedures, 339	Surgical performance: CABG,	Sleep departed Sleep deprived: no
		performed a case starting	(5%) performed by sleep-	ACC	differenc <b>ല്ട</b> n C <mark>മ</mark> BG or ACC.
RoB: low		22:00 to 05:00, or ending	deprived surgeons		d j.c
		22:00 to 07:30 and another		Assessed during surgery	com/ o
		case in the next 24-h		<u> </u>	<u> </u>
Govindarajan,	СО	Sleep deprivation: treated	NR	Surgical performance:	Sleep deprived: non-sleep deprived: no
2015		patients from midnight to		duration of surgery	difference n deration of surgery, even after
		07:00 and performed a			stratificaten by type of procedure.
RoB: low		subsequent case on the			2025 gies
		same day	N	6 1 1	. 01
Amirian, 2014	BA	17-h night call shift;	Naps pre-call: 11 (37%)	Surgical performance:	Pre- vs. post-can:
		Sleep hours during the shift:	napped for median (IQR) 90	LapSimGyn laparoscopic	1) LapSimGyn: 40 difference in total time,
RoB: high		Wrist-mounted Micro-Mini-	(58-128) min;	simulation for time, blood	blood loss, instrument path length, instrumen
		Motionlogger;	Median (IQR) sleep: 91 (62-	loss, instrument path;	angular path; nupping did not affect
		Sleepiness: KSS	123) min on the pre-call	D2 test of attention and	performance;
			night vs. 430 (329-449) on	concentration	<u>ğ</u>
					ographique de l
					que
		For peer review	v only - http://bmjopen.bmj	.com/site/about/quidelines.x	html 🙎

Risk of Bias (RoB) design   Assessment measure and time points   Assessed on pre-call and on-call day; sleepiness assessed during shift   for median (IQR) 98 (39-135) min;   Significant development of sleepiness during shift   (p<0.001), plateau score of 7 at 04:00 to 08:00.   Psychomotor performance: virtual ring transfer task for gesture-level proficiency,   Monga, 1999;   Sleep hours: self-reported (continuous)   Assessed in 3 sessions preand post-call   Assessed in 3 sessions preand	tudy	Study	Exposures or intervention		Outcomes	cted by copyrighted b
Assessed on pre-call and on- call day; sleepiness assessed during shift    Seepiness during shift   for median (IQR) 98 (39- 135) min;   Significant development of sleepiness during shift   (p<0.001), plateau score of 7 at 04:00 to 08:00.		-		Baseline		
Seep on call: 12 (40%) slept for median (10R) 98 (39- 135) min; Significant development of sleepiness during shift (p<0.001), plateau score of 7 at 0.40 to 0.80.00.  Gerdes, 2008 BA On-call shift; Fatigue differential from Fatigue: questionnaire designed by Behrenz & (units unclear); Sleep hours: self-reported (continuous)  Sleep hours: self-reported (continuous)  Shanafelt, 2010 CS Degree of fatigue as a contributor to errors (self-reported)  Shanafelt, 2010 CS Degree of fatigue as a months  Assessed for the past 3 months  Sleep hours self-reported; Firedness pre-vs. post-duty: F			time points		time points	udi
during shift    during shift   for median (IQR) 98 (39- 135) min;   Significant development of sleepiness during shift (p<0.001), plateau score of 7 at 04:00 to 08:00.			Assessed on pre-call and on-	the on-call night, p<0.001;		2) D2 test improvement in concentration,
Serior Significant development of sleepiness during shift (p<0.001), plateau score of 7 at 04:00 to 08:00.  Gerdes, 2008 BA On-call shift; Fatigue: questionnaire designed by Behrenz & (units unclear); General (range): 1-7 (units unclear); General (range): 1-8 (continuous) (cont			call day; sleepiness assessed	Sleep on-call: 12 (40%) slept	Assessed on pre-call and on-	p<0.05. Note that parameters
Significant development of sleepiness during shift (p<0.001), plateau score of 7 at 04:00 to 08:00.  Gerdes, 2008 BA On-call shift; Fatigue differential from Fatigue: questionnaire pre- to post-call (range): 1-7 (units unclear); Sleep during call (range): 1-7 (units unclear); Sleep during call (range): 1-8 smoothness, tool movement smoothness, tool movement smoothness, elapsed time (continuous)  Shanafelt, 2010 CS Degree of fatigue as a contributor to errors (self-reported)  Assessed for the past 3 months  Assessed for the past 3 months  Assessed for the past 3 months  Anesthesiologists*  Lederer, 2006 BA 24-h shift, on-call duty; Sleep hours and Sleep hours: self-reported; Sleep hours and sleep hours and post-call units and post-call u			during shift	for median (IQR) 98 (39-	call day	3) ns diffe இடி ஜீin laparoscopic simulation ti
sleepiness during shift (p<0.001), plateau score of 7 at 04:00 to 08:00.  Gerdes, 2008  BA  On-call shift; Fatigue: questionnaire pre- to post-call (range): 1-7 (units unclear); Gesture-level proficiency, Monga, 1999; Sleep during call (range): 1- Sleep hours: self-reported (continuous)  Shanafelt, 2010  CS  Degree of fatigue as a contributor to errors (self- reported)  Assessed for the past 3 months  Assessed for the past 3 months  Anesthesiologists*  Isleep hours and ROB: high  Non-call duty; Sleep hours and ROB: high  Sleep hours and ROB: high  Sleep hours and ROB: high  Non-call duty; Sleep hours and ROB: high  Sleep hours and ROB: high  Non-call duty; Sleep hours and ROB: high  Sleep hours and ROB: high  Non-call duty; Sleep hours and ROB: high  Non-call duty; Sleep hours and ROB: high  Non-call duty; Number of interruptions: self-reported; ROB: high  Non-call duty; Number of interruptions: Sleep hours and ROB: high  Non-call duty; Number of interruptions: Sleep hours and Number of interruptions: Slee						in those was bept during the shift vs. not.
(p<0.001), plateau score of 7 at 04:00 to 08:00.  Gerdes, 2008 BA On-call shift; Fatigue (differential from pre- to post-call (range): 1-7 virtual ring transfer task for gesture-level proficiency, except elagorization of the post-call (range): 1-8 (units unclear); gesture-level proficiency, except elagorization of the past 3 months  Assessed in 3 sessions preand post-call  Shanafelt, 2010 CS Degree of fatigue as a contributor to errors (self-reported)  Assessed for the past 3 months  Anesthesiologists*  Lederer, 2006 BA 24-h shift, on-call duty; Sleep hours and self-reported; Tiredness: VAS from 0 (low)  Fatigue differential from pre- to post-call (range): 1-7 virtual ring transfer task for gesture-level proficiency, except elagorization of the past 3 month self-reported; Tiredness: VAS from 0 (low)  Fatigue (questionnaire) pre- to post-call (range): 1-7 virtual ring transfer task for gesture-level proficiency, except elagorization of the post-call in psychometry gesture-level proficiency, except elagorization of the psychometry except lagorization of the post-call in psychometry perior increase dexport fatigue and post-call in psychometry gesture-level proficiency, except elagorization of the past 3 month self-reported in all the psychometry except lagorization of the past 3 months in the psychometry presenting and error, 6.99 degree of satigue as a profice of perceived recent major medical errors (self-reported) 2) of those reporting an error, 6.99 degree of satigue as a profice as the greatest of the past 3 months interruptions: self-reported; proficiency (p<0.05, all 1.1; psychometry performance: psychometry performance: psychometry except lagorization and psychometry proficiency (p<0.05, all 1.1; psychometry performance: psychometry performance: psychometry proficiency (p<0.05, all 1.1; psychometry performance: psychometry performance				-		ela ela
Assessed in 3 sessions preand post-call  CS Degree of fatigue as a contributor to errors (self-reported)  ROB: unclear  CS Degree of fatigue as a contributor to errors (self-reported)  Assessed for the past 3 months  Anesthesiologists*  Assessed in 3 to 04:00 to 08:00.  Tiredness: VAS from 0 (low)  ROB: high  Assessed in 3 to 04:00 to 08:00.  Fatigue differential from pre- to post-call (range): 1-7 virtual ring transfer task for gesture-level proficiency, except elagate (pime; no change in psychomographic programs) increase cogn (p<0.05, dimension). psychomographic programs increase dexport fatigue ration and post-call psychomographic programs increased export fatigue ration and post-call psychomographic programs. psychomographic programs increased export fatigue ration and post-call psychomographic programs increased export fatigue ration and post-call psychomographic programs. psychomographic programs increased export fatigue ration and post-call psychomographic programs. psychomographic programs increased export fatigue ration and psychomographic programs. psychomographic programs increased export fatigue ration psychomographic programs. psychomographic programs increased export psychomographic programs increased export psychomographic programs. psychomographic programs increased export psychomographic prog						201
Gerdes, 2008  BA  On-call shift; Fatigue: questionnaire pre- to post-call (range): 1-7 virtual ring transfer task for gesture-level proficiency, Monga, 1999; Sleep hours: self-reported (continuous)  Shanafelt, 2010  CS  Degree of fatigue as a contributor to errors (self-reported) Assessed for the past 3 months  Assessed for the past 3 months  Anesthesiologists*  Lederer, 2006  BA  24-h shift, on-call duty; Sleep hours and interruptions: self-reported; Sleep hours and interruptions: self-reported; Sleep hours and literature pre- to post-call (range): 1-7 virtual ring transfer task for gesture-level proficiency, except elad (p<0.05, decrepable) psychomodor performance: psychomodor psychomodor post-call psychomodor psyc						leni to
RoB: high designed by Behrenz & (units unclear); gesture-level proficiency, monga, 1999; Sleep during call (range): 1-7 (units unclear); gesture-level proficiency, hand movement psychomogogogors; increase cogn (p<0.05, designed by Behrenz & (units unclear); hand movement psychomogogogors; increase cogn (p<0.05, designed by Behrenz & (units unclear); hand movement psychomogogogors; increase cogn (p<0.05, designed by Behrenz & (units unclear); hand movement psychomogogogors; increase cogn (p<0.05, designed by Behrenz & (continuous) psychomogogogors; increase cogn (p<0.05, designed by Behrenz & (p<0.05, designed by						t Sch
ROB: high designed by Behrenz & (units unclear); gesture-level proficiency, Monga, 1999; Sleep during call (range): 1- hand movement psychometry psych	erdes, 2008	BA				1) Pre- to feet sall: decrease in all measures
Monga, 1999; Sleep during call (range): 1- hand movement smoothness, tool movement smoothness, tool movement smoothness, tool movement smoothness, elapsed time (p<0.05, daring storceased export fatigue ratings of creased export fatigue ratings of creased (R²=0.921 hours of steep declined (R²=0.933).  Shanafelt, 2010 CS Degree of fatigue as a contributor to errors (self-reported)  RoB: unclear reported)  Assessed for the past 3 months  Assessed for the past 3 months  Anesthesiologists  Lederer, 2006 BA 24-h shift, on-call duty; Mean±SD sleep: 4.1±1.7h; Psychomotor performance: reaction time, critical flicker fusion, response measure, Tiredness: VAS from 0 (low)  Number of interruptions: reaction time, critical flicker fusion, response measure, Tiredness: VAS from 0 (low)  Tiredness pre- vs. post-duty: peripheral awareness; reaction time (Ass): 252.8±39.3 vs.	- D. L.:-L		• .		_	
Sleep hours: self-reported (continuous)  Assessed in 3 sessions preand post-call  Shanafelt, 2010 CS Degree of fatigue as a contributor to errors (self-reported)  Assessed for the past 3 months  Anesthesiologists  Lederer, 2006 BA 24-h shift, on-call duty; Sleep hours and interruptions: Sleep hours and interruptions: self-reported; Sleep hours and interruptions: self-reported; Tiredness: VAS from 0 (low)  Shanafelt, 2010 CS Degree of fatigue as a contributor to errors (self-reported)  Assessed in 3 sessions pre-hours of sessions pre-hours of self-reported and post-call  Perceived recent major prediction interruptions: self-reported and post-call precision interruptions: self-reported precision interruptions: self-reported precision interruptions: self-reported precision interruptions: pre-vs. post-duty: peripheral awareness; reaction time (Pas): 252.8±39.3 vs.	os: nign					
(continuous)  Assessed in 3 sessions preand post-call  Shanafelt, 2010 CS Degree of fatigue as a contributor to errors (self-reported)  Assessed for the past 3 months  Assessed for the past 3 months  Anesthesiologists³  Lederer, 2006 BA 24-h shift, on-call duty; Sleep hours and self-reported; Sleep hours and self-reported; Sleep hours and self-reported; Oley interruptions: Rob. interruptions: self-reported; Oley interruptions: self-reported; Oley interruptions: reaction time, critical flicker interruptions: reaction time, critical flicker interruptions: reaction time (Ass): 252.8±39.3 vs. 480.3±58 reaction time (Ass): 252.8±39.3 vs.			-			
Assessed in 3 sessions preand post-call  Shanafelt, 2010 CS Degree of fatigue as a contributor to errors (self-reported)  Assessed for the past 3 months  Assessed for the past 3 months  Anesthesiologists³  Lederer, 2006 BA 24-h shift, on-call duty; Sleep hours and interruptions: self-reported; Sleep hours and interruptions: self-reported; Tiredness: VAS from 0 (low)  ROB: high  Assessed in 3 sessions pre-hours and sessions pre-hours of size preclined (R²=0.921) hours of size preclined (R²=0.933).  Assessed in 3 sessions pre-hours of size preclined (R²=0.933).  And post-call ones of size preclined (R²=0.933).  And post-call ones of self-reported independent of medical errors (self-reported)  Assessed for the past 3 medical errors (self-reported)  Assessed for the past 3 factor.  Assessed for the past 3 factor.  By Contributor to errors (self-reported)  Assessed for the past 3 factor.  Assessed for the past 3 factor.  By Contributor to errors (self-reported)  Assessed for the past 3 factor.  By Contributor to errors (self-reported)  Assessed for the past 3 factor.  By Contributor to errors (self-reported)  Assessed for the past 3 factor.  By Contributor to errors (self-reported)  Assessed for the past 3 factor.  By Contributor to errors (self-reported)  Assessed for the past 3 factor.  By Contributor to errors (self-reported)  Assessed for the past 3 factor.  By Contributor to errors (self-reported)  Assessed for the past 3 factor.  By Contributor to errors (self-reported)  Assessed for the past 3 factor.  By Contributor to errors (self-reported)  Assessed for the past 3 factor.  By Contributor to errors (self-reported)  Assessed for the past 3 factor.  By Contributor to errors (self-reported)  Assessed for the past 3 factor.  By Contributor to errors (self-reported)  By Contributor to errors (self-reported				SII		7 (0 7
Assessed in 3 sessions pre- and post-call  Shanafelt, 2010 CS Degree of fatigue as a contributor to errors (self- RoB: unclear reported)  Assessed for the past 3 months  Anesthesiologists³  Lederer, 2006 BA 24-h shift, on-call duty; Sleep hours and interruptions: and interruptions: self-reported; O.8±1.1; fusion, response measure, Tiredness: VAS from 0 (low)  Assessed in 3 sessions pre- and post-call and po			(continuous)		sinootiness, elapsed time	· · · · · · · · · · · · · · · · · · ·
and post-call  Shanafelt, 2010 CS Degree of fatigue as a contributor to errors (self-reported)  RoB: unclear  Assessed for the past 3 months  Anesthesiologists <sup>a</sup> Lederer, 2006 BA 24-h shift, on-call duty; Sleep hours and sold interruptions: Sleep hours and interruptions: self-reported; Sleep hours and interruptions: self-reported; O.8±1.1; fusion, response measure, interruption interruptions: self-reported; O.8±1.1; fusion, response measure, Tiredness: VAS from 0 (low)  To Degree of fatigue as a NR Perceived recent major 1) Prevalence of perceived recent remedical errors (self-medical err			Assessed in 3 sessions pre-		Assessed in 3 sessions pre-	<b>D</b> <
Shanafelt, 2010 CS Degree of fatigue as a contributor to errors (self-reported)  ROB: unclear reported)  Assessed for the past 3 months  Anesthesiologists  Lederer, 2006 BA 24-h shift, on-call duty; Sleep hours and Sleep hours self-reported; O.8±1.1; fusion, response measure, Tiredness: VAS from 0 (low)  Tiredness: VAS from 0 (low)  Tiredness pre- vs. post-duty: peripheral awareness; reaction time (Pas): 252.8±39.3 vs.			•			<u>a</u> . <u>3</u> .
RoB: unclear reported)  Assessed for the past 3 months  Anesthesiologists <sup>a</sup> Lederer, 2006  BA  24-h shift, on-call duty; Mean±SD sleep: 4.1±1.7h; Psychomotor performance: Pre- vs. post-duty, mean±SD: Sleep hours and Number of interruptions: reaction time, critical flicker interruptions: self-reported; 0.8±1.1; fusion, response measure, Tiredness: VAS from 0 (low)  Tiredness pre- vs. post-duty: peripheral awareness; reaction time (2) Of those reporting an error, 6.99 degree of sating and error, and	nanafelt, 2010	CS	· · · · · · · · · · · · · · · · · · ·	NR		1) Prevalence of perceived recent major
Assessed for the past 3 months  Anesthesiologists³  Lederer, 2006 BA 24-h shift, on-call duty; Sleep: 4.1±1.7h; Psychomotor performance: Pre- vs. post-duty, mean±SD: Sleep hours and Number of interruptions: reaction time, critical flicker 1) Psychometrical flicker 2) Pre- vs. post-duty: peripheral awareness; reaction time (Mas): 252.8±39.3 vs. 480.3±58 flicker 2) Psychometrical flicker 2) Psychometrical flicker 3) Psychometrical flicker 4) Psychometrical f			contributor to errors (self-		medical errors (self-	medical egror: \$9%;
Assessed for the past 3 months months  Anesthesiologistsa  Lederer, 2006 BA 24-h shift, on-call duty; Mean±SD sleep: 4.1±1.7h; Psychomotor performance: Pre- vs. pot-duty, mean±SD: Sleep hours and Number of interruptions: reaction time, critical flicker 1) Psychomotor performance: 1) Psychomotor performance: Pre- vs. pot-duty; mean±SD: Sleep hours and Number of interruptions: reaction time, critical flicker 1) Psychomotor performance: 1) Psychomotor performance: Pre- vs. pot-duty; mean±SD: 1) Psychomotor performance: 2) Pre- vs. pot-duty; peripheral awareness; reaction time (2) Psychomotor performance: 2) Pre- vs. pot-duty; peripheral awareness; reaction time (2) Psychomotor performance: 2) Pre- vs. pot-duty; peripheral awareness; reaction time (2) Psychomotor performance: 2) Pre- vs. pot-duty; peripheral awareness; reaction time (2) Psychomotor performance: 2) Pre- vs. pot-duty; peripheral awareness; reaction time (2) Psychomotor performance: 2) Pre- vs. pot-duty; peripheral awareness; psychomotor performance: 2) Pre- vs. pot-duty; psyc	oB: unclear		reported)		reported)	2) Of thosareparting an error, 6.9% listed
Anesthesiologistsa  Lederer, 2006 BA 24-h shift, on-call duty; Mean±SD sleep: 4.1±1.7h; Psychomotor performance: Pre- vs. pot-duty, mean±SD:  Sleep hours and Number of interruptions: reaction time, critical flicker 1) Psychometric testing: recognition interruptions: self-reported; 0.8±1.1; fusion, response measure, time (ms) 439 34±50.8 vs. 480.3±58 Tiredness: VAS from 0 (low) Tiredness pre- vs. post-duty: peripheral awareness; reaction time (48s): 252.8±39.3 vs.						degree of atigoe as the greatest contributing
Anesthesiologistsa  Lederer, 2006 BA 24-h shift, on-call duty; Mean±SD sleep: 4.1±1.7h; Psychomotor performance: Pre- vs. pot-duty, mean±SD:  Sleep hours and Number of interruptions: reaction time, critical flicker 1) Psychometric testing: recognition interruptions: self-reported; 0.8±1.1; fusion, response measure, time (ms) 439 34±50.8 vs. 480.3±58 Tiredness: VAS from 0 (low) Tiredness pre- vs. post-duty: peripheral awareness; reaction time (48s): 252.8±39.3 vs.			Assessed for the past 3		Assessed for the past 3	factor.
Lederer, 2006 BA 24-h shift, on-call duty; Mean±SD sleep: 4.1±1.7h; Psychomotor performance: Pre- vs. post-duty, mean±SD: Sleep hours and Number of interruptions: reaction time, critical flicker 1) Psychometrise testing: recognition RoB: high interruptions: self-reported; 0.8±1.1; fusion, response measure, time (ms) 1/439 34±50.8 vs. 480.3±58 Tiredness: VAS from 0 (low) Tiredness pre- vs. post-duty: peripheral awareness; reaction time (24s): 252.8±39.3 vs.			months		months	te n
Sleep hours and Number of interruptions: reaction time, critical flicker 1) Psychologie tribetesting: recognitions: RoB: high interruptions: self-reported; 0.8±1.1; fusion, response measure, time (ms) (max 139 k	nesthesiologists <sup>a</sup>					<b>T</b> 5
RoB: high interruptions: self-reported; 0.8±1.1; fusion, response measure, time (ms) 439 3±50.8 vs. 480.3±58 Tiredness: VAS from 0 (low) Tiredness pre- vs. post-duty: peripheral awareness; reaction time (4s): 252.8±39.3 vs.	ederer, 2006	BA	24-h shift, on-call duty;	Mean±SD sleep: 4.1±1.7h;	Psychomotor performance:	0 •
Tiredness: VAS from 0 (low) Tiredness pre- vs. post-duty: peripheral awareness; reaction time (4s): 252.8±39.3 vs.			·	Number of interruptions:	reaction time, critical flicker	1) Psychometric testing: recognition reaction
	oB: high			•		time (ms) 439 ±50.8 vs. 480.3±58.9; motor
to 100 (high) $20.0\pm 27$ E vs. E0. E±10.0 Concentration ability: scale total reaction time (ms): 600.0±72.						
			to 100 (high)	30.9±27.5 vs. 59.5±18.9,	Concentration ability: scale	total reaction the (ms): 690.8±73.4 vs.
Q			Assessed was said asset duty.	p=0.01.		746.5±113.7; critical flicker fusion (Hz):
			Assessed pre- and post-duty		(maximum tiredness)	29.0±2.3 vs. 28 ±3.7; response measure
Assessed pre- and post-duty  Assessed pre- and post-duty  For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml					Accessed are and nest duty	(pixels): 647.8 26.7 vs. 598.3±138.1,

Study	Study	Exposures or intervention		Outcomes	Associations between exposure and outcom
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	967
		time points		time points	periphera awaxeness task recognition time:
					58.9±59.2 <b>0</b> /s. 5 <b>0</b> /6±47.5;
					2) Concer 和 ability: 26.4±23.5 vs.
					56.3±23.0% pg 03007.
Chang, 2013	CS	15-h in-house overnight call;	Median (IQR) ESS: 9 (9),	Psychomotor performance:	1) Aftern  seline vs. pre-call: no
		Sleepiness pre-call: ESS ≥9;	64% scored ≥9;	reaction time; CCPT II; N-	differencಹ ಗ್ರಹ್ಮಚ್ಚಿction time, CCPT, N-back, of
RoB: unclear		Sleep hours: self-reported	Median (IQR) hours slept	back; HVLT (3 trials of 12	HVLT; to 18.
		(continuous)	during shift: 1 (0-3).	words)	Morning b தெரியு vs. post-call:
		Classians		Assessed at baseline and our	1) No change auditory or visual reaction
		Sleepiness assessed pre-call, sleep hours during call		Assessed at baseline and pre-	time; ne i o
		sleep nours during call		and post-call	2) CCPT (tss 连京): No change in detectability response 歌魚 tit reaction time,
					omissions nissions;
					3) N-back素 accuracy: no change for auditory
					visual, or mean N-value;
					4) HVLT (tascora): mean for trials 1-3: 48.6±7.
					vs. 41.5±9 (p 0.04); delayed recall: ns;
					5) No cor at lation between ESS scores pre-cal
					or sleep during hift and any measure of
					psychomotor performance.
Gander, 2000	CS	Nights of work-related sleep	NR	Risk of fatigue-related errors:	1) Risk of htigwe-related errors increased wi
D - D. I		disturbance: self-reported		questionnaire modelled after	increasing lights of work-related sleep
RoB: low		(continuous)		Gravenstein et al., 1990	disturbance: RP 1.25, 95% CI: 1.06-1.49.
		Assessed for the past 6		Assessed for the past 6	une (
		months		months	9, 2 5log
Saadat, 2017	CS	Sleep deprivation due to an	NR	Reaction time: PVT	Mean (SD reaction time was slower post-cal
		overnight call shift			(297.76 (83.75) vs. on a regular day (266.58
RoB: low				Assessed after an overnight	(38.35)), p=0.0 <b>2</b> 7.
				call shift and the morning of	en
				a regular (non-call) day	ence Bibliographique de l
					<b>≅</b>

					5/bmjopen-2018- cted by copyrig
Study Risk of Bias (RoB)	Study design	Exposures or intervention Assessment measure and time points	Baseline	Outcomes  Assessment measure and time points	Associations between exposure and outcon  - Sc 96  - C U 7  - O
Gander, 2008 RoB: unclear	NC	Sleep loss across consecutive working days or on-call work: Wrist- mounted Actiwatch (Mini Mitter, Bend, Oregon, US), sleep and duty diary  Assessed over a 2-week period including a weekend	≥2 hours sleep 8% of 24-h periods that included day work vs. 14% that included day + call; Sleep hours: mean 0.6h less sleep when working day shifts (p=0.014) and 0.8h less sleep when working day shifts + call (p=0.013) vs. off.	Psychomotor performance: PVT  Assessed within 2 hours preand post-call	1) In fixed model analysis for reaction time including geepytime since waking, work hou acute sleef (1) a support of the slower medion time including geepytime since waking, work hou reaction time including geepytime since waking as ociated with poorer performation times by the slowest 10%, F <sub>(1,185)</sub> =5.1 p<0.05; to provide across 12 consecutive work days: no change in pre-duty reaction times by
ER or ICU physician	s	of rostered shifts or on-call	<sup>2</sup> 664		post-duty te to times slowed linearly, median -388 6 0.001; decline in performan across 10 10 10 10 10 10 10 10 10 10 10 10 10
Sanches, 2015 RoB: high	CS	Acute sleep deprivation (<5h of night sleep after a night shift of 12h) Sleep hours: 7-day Actigraphy via SenseWear® Pro2 Armband; Sleepiness: ESS; Sleep quality: PSQI  Assessed the week and night before the psychomotor tests	Non-sleep deprived vs. sleep deprived:  PSQI >5: 0% vs. 33%, ns; ESS≥10: 11% vs. 67%  Sleep time (mean±SD) in week before tests: duration and number of naps higher in sleep deprived group, but diurnal sleep hours lower, 428.6±30.1 vs. 375.8±55.9, p=0.038;  Sleep quality (mean±SD): week before tests: 3.3±0.7 vs. 2.6±0.3, p=0.013; night before tests: 3.1±0.8 vs. 1.9±1.0, p=0.020.	Psychomotor performance via Battery Test Reaction 5 (v1): StimulTest, InstrucTest, MovemTest; TP test of visual attention  Assessed on morning after night shift 8	Sleep deprived roup vs. non-sleep deprived mean±SD 1 linstructurest: sorrect answers: 169.4 (16.0) 148.3 (28.7), p. 0.070; wrong answers: ns; perfection index (%): 99.6 (0.3) vs. 98.9 (1.3) p=0.021; perfection index (%): 99.6 (0.3) vs. 98.9 (1.3) p=0.021; perfection index (%): ns; response latency (sec/click): ns; perfection index (%): ns; response latency (sec/click): 1.0 (0.1) vs. 1.24 (0.1), p=0.022; 3) Movembest 1 s for any parameter; 4) TP: ometed symbols: 34.2±18.4 vs. 62.7±44.0 (p=0.034; concentration index (%): 13.8±8.6 (2.29.2±26.4, p=0.031; correct/wrong ymbols: ns; Correlations between sleep and tests: 1) TP for sleep bours nights 1-6: omitted symbols: r=-0.66, p=0.011 for non-sleep-
		For moon novino	v carla de la	.com/site/about/guidelines.>	raphique de l

cted by copyrigh 6/bmjopen-2018-0

Study	Study	Exposures or intervention		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and time points	Baseline	Assessment measure and time points	967 or
					deprived as foxsleep-deprived; concentration index (%) =-0.0359, p=0.037 for sleep-deprived, gs. fox p=0.037 for sleep-deprived; r=-0.359, p=0.037
Generalists <sup>b</sup>					Do tes
Harbeck, 2015	CS	24-hours on-call shift with sleep disturbance: self-reported number of sleep disturbances and hours of sleep per night  Assessed before a normal day shift, and after a 24-h on call shift	1) Sleep hours on a normal day vs. following a 24-h shift: <2 hours: 0 vs. 5.9%; 2-4 hours: 5.9% vs. 47.1%; 4-6 hours: 11.8% vs. 35.3%; >6 hours: 82.4% vs. 11.8% 2) Number of sleep disturbances a normal day vs. following a 24-h shift: 0: 82.4% vs. 11.8%; 1: 11.8% vs. 35.3%; 2: 5.9% vs. 47.1%; 3: 0% vs. 5.9%; 4: 0% vs. 0%; >4: 0% vs. 0%	Neurocognitive parameters: computerized attentional test (vigilance, alertness); D2 letter cancellation test (divided attention); Trail Making Test (visual attention, task switching); Digit Span, Digit Symbol Substitution Test, Weschler Memory Scale (memory functions)  Assessed before a normal day shift, and after a 24-h on call shift	Intrinsic a least section and vigilance with the property of t
Mixed specialties o	r undefine	ed populations			<u> </u>
Chen, 2008 RoB: high	CS	Sleepiness: ESS score ≥11  Time points NR	Mean±SD ESS score: 7.8±4.0, range: 0-20, 23% had scores ≥11.	Impact on work and personal life: Impact Questionnaire with a 5-point Likert scale from 1 (strongly agree) to 5 (strongly disagree)  Time points NR	1) Impact core correlated with ESS, r=0.31, p<0.05; (2) 2) ESS score was higher among physicians who agree/strongly gree vs. other response: written an incess rect order: 8.8 vs. 7.3, p=0.02; might fall asless while examining a patient: 13.2 vs. 7.7, p=0.001; look forward to sleeping at grand rounce 10.4 vs. 7.4, p=0.002;
		For peer revie	w only - http://bmjopen.bmj.	.com/site/about/guidelines.x	lographique de l

			BMJ Oper	n	5/bmjopen-2018 cted by copyrig
Study Risk of Bias (RoB)	Study design	Exposures or intervention Assessment measure and time points	Baseline	Outcomes Assessment measure and time points	Associations between exposure and outcome of the second outcome ou
		~O	Deer .		3) No difference in ESS score for those who agree/strongly or gree vs. other response: work is unaffected by sleep loss and fatigue, thinking is unaffected by sleep loss, sleep loss and fatigue affected by medical decisions, have heard of company medical decisions, have made medical experience post-call days, have made medical experience and fatigue;  4) Higher company of the properties of
Heponiemi, 2014 RoB: low	CS	Sleeping problems: 4-item Jenkins Scale on 6-point scale from 1 (never) to 6 (every night)  Assessed in 2006	Mean±SD (range) score: 2.30±1.00 (1-6)	Work ability: Work Ability Index on scale from 1 (could not work at all) to 10 (best work ability)  Assessed in 2010	1) On-call and hand an indirect effect on work ability (R <sup>2</sup> 0.15,95% CI: -0.122, -0.031, p<0.001) frough two mediators (work interference with family, sleeping problems); 2) Sleeping problems inversely associated with work ability, β=0.29, p<0.001.
Kanieta, 2011 RoB: unclear	CS	Sleep hours: self-reported (continuous) Sleepiness and sleep difficulties: 5-point scale from 1 (never) to 5 (always); Insomnia: ≥3 sleep difficulties  Assessed for the past month	Insufficient rest: 32.5%; Daytime sleepiness: 3.5%; Insomnia: 20.0%; Sleep time (mean±SD min): 279.8±60.9	Self-reported medical incidents: 4-point scale from 1 (never) to 4 (often)  Assessed for the past month	1) Prevalence of medical incidents (% (95% CI)): sleep tep wed (26.8% (24.2, 29.4)) vs. no (15.2% (13.7, 16.7)), p<0.01; insomnia (24.8% (21.6, 28.0)) insomnia vs. not (17.6% (16.2, 19.0)), p<0.01; 56h sleep (18.3% (16.8, 19.8)) vs. <6h (20.7% (18.8, 24.6)), p=0.03; 2) Predictors of medical incidents in multivariate model including personal and work-related factors (OR (95% CI)): lacking redue to sleep de rivation vs. not (1.65 (1.33-2.04)), p<0.01) insomnia vs. not (1.45 (1.16-1.82), p<0.01); of or sleep hours.
		For peer review	v only - http://bmjopen.bmj	.com/site/about/guidelines.x	Bibliographique

5/bmjopen-2018-0; cted by copyright

Study	Study	Exposures or intervention		Outcomes	Associations between exposure and outcome
Risk of Bias (RoB)	design	Assessment measure and	Baseline	Assessment measure and	- ncl
		time points		time points	udi or
Sexton, 2001	CS	Fatigue as a factor	NR	Performance effectiveness	1) "When the stigged, I perform effectively during
		impacting performance		measured by 1 question:	critical phases of operations/patient care":
RoB: high				agree, neutral, disagree	Anestheti 🖁 🞢 🖁 agree; 15% neutral; 38%
		Time points NR			disagree; $g$
				Time points NR	Surgical: 🏧 🌋 ree; 12% neutral; 18%
					disagree. 20
Shirom, 2006	CS	Tiredness and exhaustion:	NR	Quality of care: Adapted 15-	1) Quality ( predicted by
		SMBM Physician Fatigue		item SERVQUAL with a 5-	fatigue in a figoral incorporating several other
RoB: low		Subscale on a 7-point scale		point Likert scale from 1	compone සූ ජි ජූ burnout, β=0.17, p<0.05.
	from 1 (almost never) to 7		(very small extent) to 5 (very	No.	
	(always)		large extent)	ided	
		Time points NR		Time points NR	from ABES
Smith, 2017	CS	Sleep deprivation: self-	NR	Perceived competence: self-	Some physicians indicated that continual
		reported via open-ended		reported via open-ended	tiredness and shaustion led to concerns that
RoB: moderate		comments		comments	it would attect their competence; some felt
					that profe≝sior performance was
		Time points NR		Time points NR	compromed times of physical and mental
					fatigue. 🚡 💆
Tanti, 2017	CS	Fatique: questionnaire on	NR	Prescribing errors:	Perception of the contribution of fatigue to
		contributors to prescribing		questionnaire on	prescribing errers differed by physician type
RoB: high		errors, with a 5-point Likert		contributors to prescribing	(p<0.05): 젊% of community doctors, 96%
		scale (very high to very low		errors, with a 5-point Likert	hospital dector, 8% of office-working doctors
		association)		scale (very high to very low	perceived ver high or high association
				association)	between tigute and prescribing errors.
		Time points NR			0 , 2
				Time points NR	2025 gies.

<sup>&</sup>lt;sup>a</sup>Includes studies of anesthetists, where these were physicians.

<sup>&</sup>lt;sup>b</sup>Includes primary care physicians, internal medicine physicians, and general practitioners.

ACC: aortic cross-clamp time; BA: before-after; CABG: cardiopulmonary bypass time; CCPT II: Connor's Continuous Performance Test II; CI: configence interval; CO: cohort; CS: cross-sectional; ER: emergency; ESS: Epworth Sleepiness Scale; h: hour(s); HVLT: Hopkin's Verbal Learning Task; Hz: Hertz; ICU: intensive care unit; IQR: interquartile range; KSS: Karolinska Sleep Scale; min: minutes; ms: millisecond(s); N-back: Dual N-back test; NA: not applicable; NR: not reported; ns: not statistically significant; OR: odds ratio; PSQI: Pittsburgh Sleep Quality Index; PVT: Psychomotor vigilance Performance Task; RR: risk ratio; RCT: randomized controlled trial; RoB: Risk of Bias D: standard deviation; SE: standard error; SERVQUAL: Service Quality Measure; SMBM: Shirom-Melamed Burnout Measure; TP: Toulouse-Piéron test; TS: time series; US: pitted States of America; vs.: versus

Patient outcomes related to fatigue or sleen restriction among physicians in independent practice

Study	Study	Exposures		Outcome Measures	Associations that tween exposure and outcom
Risk of Bias (RoB)	design	Intervention or assessment	Baseline	Assessment scale and time	on 2,
		scale and time points		points	າ <u>ຍ</u> 21
Surgeons					or Se
Chu, 2011	СО	Sleep deprivation: moderate	Of 4,047 procedures, 83	Chart review: mortality, surgical	1) 0-3 vs 🖟 🧃 😸. >6 hours of sleep: No
		(3-6 h) or severe (<3-h) sleep	(2.1%) performed by	complications, length of stay	difference is is incidence of mortality, incidence
RoB: low		deprivation the night before	severely sleep-deprived,		of 10 ma
		surgery (self-reported hours)	1,595 (39.4%) by	Assessed during and post-	3.6% vs. 🕳 🏂 . 0.8%, p=0.03), ICU length of
			moderately sleep-	surgery	stay; in-hថ្ងៃទៀវជា length of stay (days): 7.0 vs.
			deprived surgeons		6.0 vs. 7. <b>5 2</b> 001.
Ellman, 2004	CO	Sleep deprivation: performed	Of 6,751 procedures,	Chart review: mortality, surgical	1) Sleep 🔐 🚾 d vs. non-sleep deprived: no
		a case starting 22:00 to 05:00,	339 (5%) were	complications, length of stay	difference i திரைப்பு, need for blood
RoB: low		or ending 22:00 to 07:30 and	performed by sleep		productsနှင့်ကြွှာlications (operative,
		performed a subsequent case	deprived surgeons	Assessed during and post-	neurologi, de Hal, infectious, pulmonary), in-
		in the next 24-h		surgery	hospital <b>ទ្ធា</b> អ្នក of stay.
Govindarajan,	CO	Sleep deprivation: treated	NR	Chart review: mortality, surgical	1) Sleep prized vs. non-sleep deprived: no
2015		patients from midnight to		complications, readmission,	difference in mortality, surgical complications
		07:00 and performed a		length of stay	readmiss ns ithin 30 days, or length of stay
RoB: low		subsequent case on the same			ining,
		day		Assessed during and post-	ე <mark>. ლ</mark>
				surgery	a b
Rothschild, 2009	CO	Sleep deprivation: daytime	NR	Chart review: frequency of	1) Post-ng httme vs. control: no difference in
		procedures following an		adverse surgical complications	number <b>a</b> prosedures with complications,
RoB: low		overnight procedure;			total nunguer of complications, preventable
		Sleep opportunity: 0-6h, <6h		Assessed during and post-	complications type of complications;
				surgery	2) Operaging room procedures with
					complica On (95% CI): 8.5% for 0-6h
					sleep vs. 4.1% or >6h sleep, 2.70 (1.13-6.48),
					p=0.03; <b>½</b> 25
					3) All procedures with complications, OR (95%
					CI): 6.2% for (66h sleep vs. 3.4% for >6h sleep
					1.72 (1.02-2.8 <b>岁</b> ), p=0.04.
					<u></u>
					<u> </u>
					Bibliographique de l
					aph
					niqu
				mj.com/site/about/guidelines.xl	<u> </u>

Page 80 of 86

			BMJ Open									
					5/bmjopen-2018- cted by copyrig							
Study	Study	Exposures		Outcome Measures	Associations between exposure and outcom							
Risk of Bias (RoB)	design	Intervention or assessment	Baseline	Assessment scale and time	- 967							
		scale and time points		points	udi							
Schieman, 2007	СО	Fatigue: surgeon billed for	Of 270 procedures, 22	Chart review: surgical	1) Fatigued vsxnon-fatigued surgeons: no							
		clinical work after 22:00 the	(8%) were performed by	complications, length of stay,	differenc <b>्र</b> in i <b>क्र</b> ra- or post-operative							
RoB: low		night before surgery	fatigued surgeons	mortality, cancer recurrence	complica இடி இte, length of stay, in-hospital length of stay இancer recurrence.							
				Assessed during and post-	r <u>eig</u>							
				surgery	r 20 ner ate							
/inden, 2014	СО	Sleep deprivation (at risk):	Of 94,183 surgeries,	Chart review: conversion to	1) At risk x net at risk surgeon: no difference							
		surgeon worked 00:00 to	2,078 (2.2%) were	open procedure (from	in incide							
RoB: low		07:00 and performed surgery	performed by surgeons	laparoscopic), iatrogenic	iatrogeni <b>ट्र</b> िंम्ब्रुडॉंes, mortality, in either							
		07:00 to 18:00	who were 'at risk'	injuries, mortality	univariate analyses.							
				Assessed during and post-	ed f							
				surgery	mi Bron							
Obstetricians					ng h							
Rothschild, 2009	СО	Sleep deprivation: daytime	NR	Chart review: frequency of	1) Post-nighttine vs. control: no difference i							
		procedures following an		adverse obstetric complications	number र्द्ध præedures with complications,							
RoB: low		overnight procedure;			total con ಕ್ರೀ licಹions, preventable							
		Sleep opportunity: 0-6h, <6h		Assessed during and post-	complications type of complications;							
				delivery	2) No assaciation between sleep deprivation							
					and propartion of procedures with							
					complica onsons complica on one complica on one complica on one complication of the complex complex of the comp							
					of sleep 📅 portunity.							
: confidence interv	al; CO: col	hort; h: hours; NR: not reported; (	DR: odds ratio; RoB: Risk of E	delivery  Bias; SD: standard deviation; US: Uni	and propartical of procedures with complication of procedures with complication of sleep grounity.  Ited States of America; vs.: versus  of sleep grounity.  ted States of America; vs.: versus  ogic 9,							
					25 at Agence Bibliographique de l							

## **Dichotomous outcomes**

Outcome or subgroup	Number	Number of	Pooled risk	Heterogeneity	
	of studies	participants	ratio (95% CI)	P	l <sup>2</sup>
1.1 Patient mortality	5	60,436	0.98 (0.84, 1.15)	0.73	0%
1.2 Intra-operative	3	19,798	suppressed	0.007	82%
complications					
1.2.1 Surgical procedure	3 <sup>a</sup>	14,896	suppressed	<0.001	88%
1.2.2 Obstetric procedure	1 <sup>a</sup>	4,902	suppressed	NA	NA
1.3 Post-operative	5	60,201	0.99 (0.95, 1.03)	0.45	0%
complications					

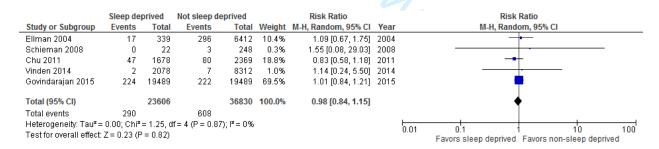
<sup>&</sup>lt;sup>a</sup> Rothschild, 2009 is represented in both analyses

## **Continuous outcomes**

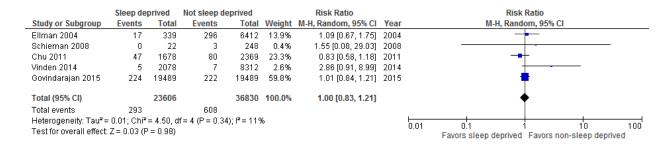
Outcome or subgroup	Number of	Number of	Pooled mean	Heterogeneity	
	studies	participants	difference (95% CI)	Р	l <sup>2</sup>
1.4 Operating time (minutes)	4	50,046	-0.14 (-1.60, 1.33)	0.70	0%
1.5 Length of hospital stay (days)	4	50,046	suppressed	<0.001	86%
1.5.1 Cardiac surgeries	2	10,798	suppressed	0.01	84%
1.5.2 Elective surgeries	1	38,978	suppressed	NA	NA
1.5.3 Anterior resection for anal	1	270	suppressed	NA	NA
cancer		<b>V</b> ,			

CPBT: cardiopulmonary bypass time; NA: not applicable

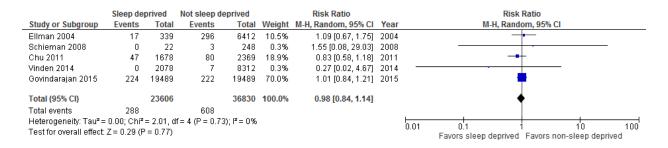
## 1.1 Patient mortality



#### Sensitivity analysis using highest possible number of events for Vinden 2014



## Sensitivity analysis using lowest possible number of events for Vinden 2014



# 1.3 Post-operative complications

	Sleep de	prived	Not sleep d	eprived		Risk Ratio		Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI		M-H, Random, 95% CI
Chu 2011	259	1678	404	2369	7.8%	0.91 [0.78, 1.04]		
Ellman 2004	29	339	480	6412	1.2%	1.14 [0.80, 1.64]		<del></del>
Govindarajan 2015	3527	19489	3543	19489	89.4%	1.00 [0.95, 1.04]		
Schieman 2008	12	22	164	248	1.0%	0.82 [0.56, 1.22]		<del></del>
Vinden 2014	14	2031	72	8124	0.5%	0.78 [0.44, 1.38]		<del></del>
Total (95% CI)		23559		36642	100.0%	0.99 [0.95, 1.03]		•
Total events	3841		4663					
Heterogeneity: Tau <sup>2</sup> =	0.00; Chi²	= 3.69, d	f = 4 (P = 0.4)	5); I² = 0%			<del>_</del>	<u>-</u>
Test for overall effect:	•						0.2	0.5 1 2 5 Favors sleep deprived Favors non-sleep deprived

## 1.4 Operating time (minutes)

1.4 Operating time (minutes)  Sleep deprived Not sleep deprived Mean Difference Mean Difference										
	Slee	ep depriv	ed	Not sl	eep depr	ived		Mean Difference		Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	Year	IV, Random, 95% CI
Ellman 2004	107.7	55.24	339	107.4	56.05	6412	5.9%	0.30 [-5.74, 6.34]	2004	+
Schieman 2008	114	226.6	22	148	226.6	248	0.0%	-34.00 [-132.80, 64.80]	2008	<del></del>
Chu 2011	110.2	119.13	1678	114	120.88	2369	3.8%	-3.80 [-11.30, 3.70]	2011	<u>-+</u>
Govindarajan 2015	156	75.56	19489	156	80	19489	90.2%	0.00 [-1.54, 1.54]	2015	<b>.</b>
Total (95% CI)			21528				100.0%	-0.14 [-1.60, 1.33]		
- '	Heterogeneity: Tau* = 0.00; Chi* = 1.42, df = 3 (P = 0.70); i* = 0%  Test for overall effect: Z = 0.18 (P = 0.86)  Test for overall effect: Z = 0.18 (P = 0.86)  Test for overall effect: Z = 0.18 (P = 0.86)									



Open cted b	Page 84 o
octed by copyrig	
-021967 ght, inclu	Reported on page #
on 2	
alysis, or both.	1
prer En	
ble: background; objectives; data sources; strate gibility criteria, synthesis methods; results; limitations; conductions of a gibility criteria, and implications of a gibility criteria, and	2 of
Downt Su tex	
of what is already known.	4-5
dressed with reference to participants, interventions, comparisons,	5
BES.	
an be accessed (e.g., Web address), and, if a labe, provide registrati	on 5
follow-up) and report characteristics (e.g., years can sidered, language, ing rationale.	6
th dates of coverage, contact with study authors to identify additional	5-6
ne database, including any limits used, such that incould be repeated.	Supplementar file 1
ng, eligibility, included in systematic review, and, is applicable, included	l in 6-7
g., piloted forms, independently, in duplicated any processes for	7
ıght (e.g., PICOS, funding sources) and any assum ដល់លេខ and simplificat	ions 7
individual studies (including specification of whether this was done at t s to be used in any data synthesis.	the 7
io, difference in means).	8
ing results of studies, if done, including measures وأقوة consistency (e.g., ا	12) 8
ing	



age 85 of 86		BMJ Open d t	
Appendix 1. PRISM	MA checl	ppen-2018	
Section/topic	#	Checklist item 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).	8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression	8
RESULTS		e ign	
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reason exclusions at each stage, ideally with a flow diagram.	8, Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follows approvide the citations.	8-11, Table 1, Supplementary file 2
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see it by a larger of the larger of	11, Supplementary file 3
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary date for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	p. 12-18; Supplementary file 4; figures 2-4
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consecutive.	p. 12-18, figures 2-4
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Not applicable
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Supplementary file 5
DISCUSSION		olog	
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider the relevance to key groups (e.g., healthcare providers, users, and policy makers).	18-19
, 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).	19-20
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for Exture research.	20
FUNDING		- Ogr	
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.	21
5		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	

\*\*September 2018. Downtoaded Home No. 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyss. Including for uses Related to use and data mining. At training, and similar technologies.

...d from http://bmjopen.bmj.com/
ur (ABES)
...data mining, Al training, and similar te.