Health Psychology Review

ISSN: 1743-7199 (Print) 1743-7202 (Online) Journal homepage: www.tandfonline.com/journals/rhpr20

Systematic review of the effects of decision fatigue in healthcare professionals on medical decision-making

Mona Maier, Daniel Powell, Peter Murchie & Julia L. Allan

To cite this article: Mona Maier, Daniel Powell, Peter Murchie & Julia L. Allan (01 Jul 2025): Systematic review of the effects of decision fatigue in healthcare professionals on medical decision-making, Health Psychology Review, DOI: 10.1080/17437199.2025.2513916

To link to this article: https://doi.org/10.1080/17437199.2025.2513916

Published online: 01 Jul 2025.

HEALTH PSYCHOLOGY REVIEW
https://doi.org/10.1080/17437199.2025.2513916

Routledge
Taylor & Francis Group

REVIEW ARTICLE

Systematic review of the effects of decision fatigue in healthcare professionals on medical decision-making

Mona Maier [a], Daniel Powell[a], Peter Murchie[b] and Julia L. Allan[c]

[a]Health Psychology, Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK; [b]Academic Primary Care, Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, UK; [c]Division of Psychology, University of Stirling, Stirling, UK

ABSTRACT

Decision fatigue is the tendency towards making less effortful decisions as the cumulative mental burden of effortful decision-making increases. Health professionals working long shifts may be particularly vulnerable to decision fatigue. This preregistered systematic review (Prospero ID = CRD42021260081, no external funding) aims to synthesise the empirical evidence on decision fatigue in the healthcare context. Systematic searches across eight databases identified 14,740 records. *N* = 82 studies (72 quantitative, 1 qualitative, 1 review, 8 expert discussions) met the inclusion criteria (health professionals/trainees; medical decisions over time; healthcare context; any design). Study quality was assessed with the MMAT or relevant JBI checklist. Narrative synthesis revealed that 45% of cases that quantitatively assessed the decision fatigue hypothesis provided evidence of significant decision fatigue effects across diagnostic, test ordering, prescribing, and therapeutic decisions. Expert discussions confirmed healthcare professionals' recognition of decision fatigue as an important phenomenon. However, decision fatigue was inconsistently defined and inadequately operationalised, reflecting limitations in current theoretical understanding of the phenomenon. To address this, we propose a new definition for greater conceptual clarity and more consistent operationalisation in future research.' Future studies should prioritise the development and testing of different theoretical explanations for decision fatigue to improve understanding and facilitate the development of appropriate interventions.

ARTICLE HISTORY
Received 25 June 2024
Accepted 27 May 2025

KEYWORDS
Decision fatigue; healthcare professionals; medical decision making; systematic review; time on task; cognitive load

Introduction

Decision-making is a critical aspect of healthcare, and over the course of a workday, healthcare professionals make a series of decisions about patient care that can, directly and indirectly, impact patient outcomes (Croskerry, 2017). Important decisions, such as formulating correct diagnoses and selecting effective and safe treatment options, are central to patient outcomes and the appropriate allocation of limited resources within the healthcare system. While scientific advancements and the application of evidence-based medicine have led to wide-reaching improvements in modern healthcare, human decision-making, including that of healthcare professionals, remains susceptible to bias (Kahneman, 2003), and suboptimal decisions are made relatively frequently (Graber, 2005). One potential driver of suboptimal care decisions is decision fatigue: a tendency towards

CONTACT Mona Maier m.maier.20@abdn.ac.uk
Supplemental data for this article can be accessed online at https://doi.org/10.1080/17437199.2025.2513916

making less effortful decisions as the cumulative burden of decision-making increases (Oakes & Patel, 2021).

While descriptions of decision fatigue in the literature vary, the term is generally used to reflect a gradual shift in decision-making that emerges over time as more decisions are made. This systematic review focuses specifically on studies that align with this understanding. Research that conceptualises decision fatigue in fundamentally different ways, such as studies examining general mental exhaustion or other forms of fatigue unrelated to sequential decision-making, is beyond the scope of this review. Decision fatigue has been observed across a diverse range of contexts where decisions are made sequentially over a sustained period. For example, credit officers in the finance sector are less likely to approve credit loans during midday compared to early in the workday (Baer & Schnall, 2021), analysts become less accurate in their forecasting as the number of forecasts they issue in a day increases (Hirshleifer et al., 2019); and academic journal editors' rejection rates are higher when reviewing larger numbers of papers (Stewart et al., 2012). One of the most commonly cited observations of decision fatigue in the literature came from an observational study showing that court judges become progressively more likely to deny parole requests (in favour of keeping the prisoner incarcerated) as court sessions wear on (Danziger et al., 2011). While this study stimulated wider research interest in the phenomenon, the analysis showed an overly large effect, which is likely to be at least partly attributable to confounds within the court data, such as the predictable scheduling of cases of different levels of severity (Glöckner, 2016; Weinshall-Margel & Shapard, 2011). However, numerous studies have since demonstrated patterns in decision-making that are consistent with decision fatigue, including studies in healthcare professionals indicating that time on task reliably influences decision-making (Oakes & Patel, 2021). For example, triage nurses working at a medical telephone helpline became increasingly likely to make conservative triage decisions as the time since their last break increased (Allan et al., 2019)General practitioners deliver fewer flu vaccinations (Kim et al., 2018) and become less likely to order cancer screenings for their patients over the course of a clinic session (Hsiang et al., 2019); and surgeons are less likely to decide that patients need surgery towards the end of shifts (Persson et al., 2019).

In healthcare settings, where clinicians regularly work lengthy periods without a break, decision fatigue may be particularly pronounced as it is hypothesised to develop in response to cumulative effort over time. For example, general practitioners working in primary care in the UK are currently seeing, and making decisions about, an average of 37 patients per day, which significantly exceeds the recommended number of encounters that are deemed a safe level ('not more than 25') by the British Medical Association (Mehlmann-Wicks, 2022). Similarly, recent surveys show that the majority of general practitioners in the UK (77%) take no breaks during 4-hour clinic sessions and that more than a quarter (28%) work full days without a proper break (Bostock, 2016). In the secondary care setting, individual patients have, on average, 13 decisions made about them during single encounters (Ofstad et al., 2018), which translates to large numbers of decisions accumulating over a healthcare professional's shift as they encounter multiple patients. Across the healthcare sector, the potential for decision fatigue to develop is high and likely increasing, as workloads grow due to an ageing population, increasingly complex health problems, and a shrinking workforce (Department of Health, 2012; Fisher et al., 2017).

Despite its potential to inform theorising about fatigue within health psychology and its potential to act as an actionable target in addressing inappropriate care decisions or to act as a marker of workload burden, decision fatigue remains an emerging and poorly understood concept. Pignatiello et al. (2020) first conducted a conceptual analysis of decision fatigue in 2020. They found that the existing literature failed to adequately describe the consequences of decision fatigue and noted a lack of robust empirical evidence of decision fatigue, particularly in the context of healthcare. To date, there has been no systematic review of the decision fatigue literature in general or in a specific field such as healthcare. The present systematic review aims to identify, evaluate, and synthesise the evidence regarding decision fatigue in healthcare professionals and its impact on medical decision-making, and to provide an overview of how decision fatigue is conceptualised theoretically in the current literature.

Review questions

For which types of medical decision-making has decision fatigue been (directly or indirectly) observed in healthcare professionals?

How is decision fatigue operationalised and measured in quantitative studies of decision fatigue in healthcare professionals?

How does decision fatigue in healthcare professionals impact medical decision-making and related outcomes?

How is decision fatigue conceptualised theoretically in the current healthcare literature?

Methods

Search strategy and study selection

The protocol for this systematic review was preregistered on PROSPERO and can be accessed at https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021260081 (Maier et al., 2021). The following electronic bibliographic databases were searched: CINAHL, Cochrane Library, Embase, OpenGrey, PubMed (incl. MEDLINE), PsycNet (PsycINFO), Scopus, and Web of Science. The search strategy aimed to find published articles of any study design, published in peer-reviewed journals in the English language. It was guided by the PECO framework (Population, Exposure, Comparator, and Outcome) (Morgan et al., 2018), with the exception of the 'comparator' component, which did not apply to this review.

The search strategy used an expansive list of synonyms and Boolean connectors for all three PECO elements: Population (healthcare professionals), Exposure (decision fatigue), and Outcome (medical decisions) to identify relevant studies. Search terms were restricted to title, abstract, and keywords. The search terms were initially developed for MEDLINE (Ovid) and then adapted for other databases. The full search strategy for one database (PubMed) can be found in Supplementary Materials, S1. Initial searches were conducted in September 2021, databases were searched from their inception, and the search was updated in January 2024. To identify additional relevant papers, manual citation searches of included papers were conducted.

After deduplication, all identified records were uploaded to the Rayyan web app for systematic reviews (Ouzzani et al., 2016). All titles and abstracts were screened for eligibility by one reviewer, MM, and 10% were independently screened by a second reviewer, JA. Full texts were then independently screened by two reviewers (MM and JA) and coded as eligible or excluded. Any disagreements were resolved by discussion.

Study inclusion and exclusion criteria

The review considered studies of healthcare professionals, trainees, or students, making decisions in any healthcare setting, including hypothetical healthcare scenarios designed for study purposes. The population scope was defined by the MeSH terms Health Personnel (Unique ID: D006282; excluding Veterinarians), and Students, Health Occupations (Unique ID: D013336). We included quantitative, qualitative, and mixed methods studies as well as opinion, narrative overview, and editorial articles, and literature reviews. No limitation was placed on study design, which meant that both observational and experimental study designs were included. Where literature reviews did not deliver any additional insights to the studies already included in our systematic review, any papers cited in the review reporting primary data were included, and the review itself was excluded.

To date there is not a single agreed gold standard measure or proxy measure of decision fatigue, therefore quantitative studies were considered eligible if they evaluated clinical decisions made under conditions where decision fatigue would be expected to emerge (i.e., where multiple decisions were made over a defined time period) and where changes related to this fatigue could be detected (i.e., where decision outcomes were reported according to time or sequence within the work period). The order number of decisions made within a work shift, time (e.g., hours) into the work shift, or time-of-day were accepted as decision fatigue surrogates, indicating periods of effortful decision-making. Comparisons between lengthy periods (e.g., between morning and afternoon, between day and night, or between different periods of 3 hours or more) were excluded because temporal trends could not be observed as they occurred. Studies using the decision fatigue scale (developed and validated by Hickman et al. (2018)) were included. The following outcome measures were included: observable clinical decision outcomes such as diagnoses (or diagnostic errors), selected/prescribed treatments, or guideline (non-)compliance, as well as observable decision-making process factors such as duration.

Qualitative studies and opinion, narrative overview, and editorial articles were considered if they specifically used the term decision fatigue, if the authors' definition broadly corresponded with the definition proposed by Pignatiello et al. (2020) in their conceptualisation of decision fatigue, and if they investigated or discussed decision fatigue consequences in terms of changes to the decision-making process, such as perceived effortfulness, or decision-making outcomes.

Data extraction and coding

For all eligible studies, one reviewer (MM) extracted data and a second reviewer (DP) independently extracted data from 20% of the sample. Before this, extraction sheets were piloted with two papers to finalise items and ensure mutual understanding of items and level-of-detail to extract. The full data extraction sheet used is found in Supplementary Materials, S2. Any discrepancies in data extraction were resolved through discussion, with consultation of a third reviewer (JA) where needed.

To answer our first research question, medical decisions (or observable decision outcomes) for which decision fatigue effects were investigated (in the quantitative papers) were categorised according to a revised version of the Decision Identification and Classification Taxonomy for Use in Medicine (DICTUM) (Ofstad et al., 2016). DICTUM classifies medical decisions into the following 10 categories: (1) Gathering additional information, (2) Evaluating test results, (3) Defining problem, (4) Drug related, (5) Therapeutic procedure related, (6) Legal and insurance related, (7) Contact related, (8) Advice and precaution, (9) Treatment goal, and (10) Deferment. This coding scheme was developed through a content-driven iterative process based on video recordings from patient-physician encounters within the hospital setting. As this context is narrower in scope than studies included in this systematic review, two minor amendments to the taxonomy were necessary to code all papers. During the review process, while coding the included papers, we identified decisions that did not fit within the existing DICTUM categories. To ensure comprehensive classification, we expanded the taxonomy by introducing two additional categories: (11) Documentation related and (12) Infection prevention related. Category 11 includes all decisions made by healthcare professionals to document medically relevant information, such as entries into the electronic health record and writing of reports. Category 12 covers decisions to take action to prevent patients and healthcare workers from being harmed by avoidable infections. Furthermore, a coding decision was made to code triage decisions under category 7 (Contact related). For brevity, in this systematic review, we abbreviate DICTUM category labels as DC1 to DC12.

Assessment of methodological quality

Eligible studies were critically appraised by one reviewer (MM). The quality of qualitative and quantitative studies was assessed using the Mixed Methods Appraisal Tool (Hong et al., 2018), opinion, narrative overview, and editorial articles were assessed using the JBI Critical Appraisal Checklist for Text and Opinion Papers (*JBI Manual for Evidence Synthesis*, 2020), and reviews were assessed using the JBI Review Checklist (*JBI Manual for Evidence Synthesis*, 2020). Quality scores were only used for guidance purposes, and no studies were excluded based on high risk of bias scores.

Data synthesis and integration

The evidence was synthesised narratively. Meta-analysis was not possible due to the inclusion of non-quantitative study designs and the heterogeneity of data and methodologies in the quantitative studies. The narrative synthesis was reported according to the PRISMA guidelines (Page et al., 2021) (see Supplementary Materials, S3 for the PRISMA Checklist).

Results

Study characteristics

The PRISMA flow chart and the full study selection process are shown in Figure 1. Database searching resulted in 14,740 records. After removing duplicates, 10,211 records remained. Screening of titles and abstracts resulted in 9,728 records being excluded leaving $n = 483$. Of the 477 full texts reviewed (6 of 483 could not be retrieved) 75 were deemed eligible for the review. A further 7 eligible full texts were identified from the reference sections of these papers, meaning the inclusion of a total of 82 studies.

Of the 82 studies, 72 were quantitative, 8 were opinion, editorials, book chapters, or narrative overviews, 1 was qualitative, and 1 was a literature review. All studies were published between 1977 and 2023, with only 3 published before 2000. Characteristics of the included studies (including study design, clinical setting, healthcare and patient sample sizes, geographic context, and others) are further described below and detailed in Table 1.

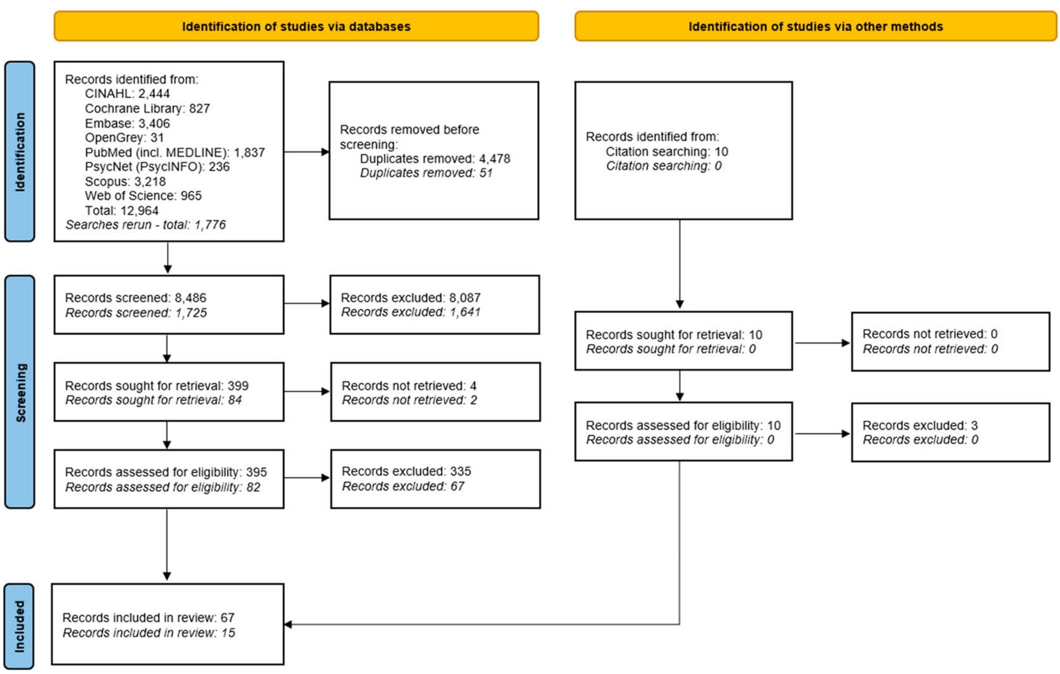

Figure 1. PRISMA 2020 flow diagram for new systematic reviews which included searches of databases.

Table 1. Characteristics of included studies (ordered alphabetically and by study type).

Study (First author & Year)	Study Design	Sample (Type of HCPs)	Setting	Country	N (HCP & Patients/ cases)	Decision Type (DICTUM) Not applicable for attributes linked to the decision-making process	Specific Decision(s) *or* Attribute(s) linked to the decision-making process	DF measure(s)
Quantitative studies								
Addis (2020)	Retrospective observational study	Anaesthesiologists, Surgeons	Hospital, Cardiology	USA	HCPs: NR Patients: 1,220	5 – Therapeutic procedure related	Blood product transfusion during cardiac surgery Red blood cell transfusion Transfusion for plasma Transfusion for cryoprecipitate Transfusion for platelets	Time of day
Agarwal (2022)	Retrospective observational study	Physicians, Residents	Hospital, Emergency	USA	HPCs: NR Patients: 16,115	4 – Drug related	Opioid prescribing	Time into shift
Allan (2019)	Observational, within-subject, repeated-measures study	Nurses	Medical telephone helpline	UK	HCPs: 150 Patients: 3,948	7 – Contact related	Triage decisions (referral within 12 hr vs outside 12-hr window)	Number of decisions made previously Time into shift Time since break
Almadi (2015)	Retrospective cohort study	Gastroenterologists, Colorectal surgeons	Hospital, Endoscopy service	Canada	HCPs: 12 Patients: 430	2 – Evaluating test result	Adenoma detection (detection of at least one adenoma on screening colonoscopy)	Time into shift Number of decisions made previously
Anastasian (2014)	Observational, retrospective study	Anaesthesiologists	Hospital, Anaesthesiology	USA	HCPs: NR Patients: 289	5 – Therapeutic procedure related	Delay extubating at the end of multi-level prone spine surgery	Time of day
Backmann (2021b)	Retrospective, observational, multicentre study	Radiologists	Radiology – population-based screening programme	Norway	HCPs: 148 Patients: 610,104 (2,937,312 readings)	2 – Evaluating test result	Mammography reading decision (true positive, false positive, true negative, false negative)	Time of day
Backmann (2021a)	Retrospective, observational, multicentre study	Radiologists	Radiology – population-based screening programme	Norway	HCPs: 148 Patients: 610,104 (2,937,312 readings)	2 – Evaluating test result	Mammography reading decision (true positive, false positive, true negative, false negative)	Order number of decision
Beg (2021)	Prospective evaluation study	Physicians, Nurses	Hospital, referral centre	UK	HCPs: 32 Patients: 196 (six historical cases per reader)	2 – Evaluating test result Not applicable*	Capsule endoscopy (CE) reading accuracy CE reading time Fatigue score* Psychomotor vigilance*	Order number of decision
Bernstein (2022)	Retrospective comparative time series study	Radiologists	Hospital, Radiology	USA	HCPs: 18 Patients: 97,671	2 – Evaluating test result	Mammography recall False-positive True-positive	Time of day

Bersani (2020)	Cluster-randomized stepped wedge trial	Physicians, Nurses, Physician Assistants, Medical students, Administrative Staff, Pharmacists	Hospital, Neurology, Oncology, General medicine services	USA	HCPs: 413 Patients: NR (8,302 logins)	1 – Gathering additional information	Use of a Patient Safety Dashboard (within Electronic Health Record)	Time of day
Bonafide (2017)	Prospective cohort study	Nurses	Hospital, general paediatrics and medically complex services	USA	HCPs: 38 Patients: 100 (3280 alarms)	2 – Evaluating test results	Response to physiologic monitor alarms	Time into shift
Canfield (2020)	Retrospective chart review	Physicians, Nurses, Pharmacists, Advanced practice providers	Hospital, Emergency Departments	USA	HCPs: NR Patients: NR (21,783,373 individual orders)	11 – Documentation related	Patient order entry (of procedures or medication) into electronic health records (right vs wrong)	Time of day
Castrén (2015)	Registry based case-control study with prospectively collected data	EMCC dispatchers and EMS personnel (including nurses, assistant nurses, and 'people with other backgrounds')	Emergency Medical Communication Center (EMCC) and Emergency Medical Service (EMS)	Sweden	HCPs: NR Patients: 986	7 – Contact related	Pre-hospital assessment decision (non-specific complaint vs any other condition)	Time of day
Chan (2009)	Prospective study	Physicians, Nurses, Gastroenterology trainees	Hospital, Gastroenterology	USA	HCPs: 22 Patients: 477	2 – Evaluating test result	Polyp identification and polyp removal Histologically confirmed polyps removed (proportion of adenomatous versus hyperplastic polyps removed)	Time of day
Changolkar (2020)	Prospective study	Physicians	Primary Care	USA	HCPs: 118 Patients: 127,021	5 – Therapeutic procedure related	Influenza vaccination	Time of day
Cheng (2011)	Pilot trial	Nurses, nursing officers, physiotherapists, and healthcare assistants	Hospital, Neurosurgical intensive care unit	China	HCPs: 21 Patients: NR (16,327 hand hygiene opportunities)	12 – Infection prevention related	Hand hygiene compliance	Time of day
Christensen (1977)	NR – Experimental study	Radiology residents	Hospital, Radiology	USA	HCPs: 14 Patients: 25 images of fictitious patients	2 – Evaluating test result	Detection of pulmonary nodules on radiographs Number of false positive detections	Time into shift
Dai (2015)	NR – Longitudinal field observation study	Nurses, Physicians, Patient care technicians, Therapists, Clinical directors, Infection preventionists, others (not defined further)	Hospital	USA	HCPs: 4157 Patients: NR (13,772,022 hand hygiene opportunities)	12 – Infection prevention related	Hand hygiene compliance	Time into shift Hours off since previous shift
Delaney (2005)	Retrospective study	Staff members of psychiatric hospital (not further specified)	Hospital, Psychiatry	USA	HCPs: NR Patients: 100 (120 restraint incidents)	5 – Therapeutic procedure related	Restraining a patient	Time of day
Dexheimer (2017)	Retrospective study	Nurses, Physicians, Anaesthesiologists	Children's hospital	USA	HCPs: 4575 Patients: NR (562,799 medication order-related alerts)	4 – Drug related	Response to medication order alert (whether the prescriber took any corrective action on order that generated alert)	Time of day

(*Continued*)

Table 1. Continued.

Quantitative studies

Study (First author & Year)	Study Design	Sample (Type of HCPs)	Setting	Country	N (HCP & Patients/ cases)	Decision Type (DICTUM) Not applicable for attributes linked to the decision-making process	Specific Decision(s) *or* Attribute(s) linked to the decision-making process	DF measure(s)
Dominello (2015)	NR – Retrospective study	Multiple members of a team (not specified), including physicians, nurses, physicists, dosimetrists, therapists, and hospital administrators	Hospital, Department of Radiation Oncology	USA	HCPs: NR Patients: 5786	5 – Therapeutic procedure related	Radiation treatment delivery errors (deviation in dose or dose distribution as delivered to the patient compared with intended dose or dose distribution)	Time of day
Edwards (2003)	Retrospective chart audit	Nurses	Medical ward in Paediatric Hospital	Australia	HCPs: NR Patients: 67	3 – Defining problem 4 – Drug related	Temperature recording Antipyretic administration	Time of day
Ergün-Sahin (2022)	Empirical investigation	Physicians	Hospital, internal medicine outpatient unit	Turkey	HCPs: 6 Patients: 37,857	1 – Gathering additional information	Diagnostic test ordering Lab test ordering Imaging test ordering	Number of decisions made previously
Evans (1984)	NR – Retrospective study	Physicians, Aneasthesiologists, etc. (not further specified)	Hospital, Maternity	USA	HCPs: NR Patients: 11,011	5 – Therapeutic procedure related	Delivery via caesarean section (vs vaginal delivery)	Time of day
Faulk (2013)	Retrospective chart review	*Not specified*	Inpatient rehabilitation facility	USA	HCPs: NR Patients: 2,282	7 – Contact related	Admission to inpatient rehabilitation facility	Time of day
Freedman (2011)	Retrospective chart review	Gastroenterologists	Gastroenterology practice	USA	HCPs: 3 Patients: 1,486	2 – Evaluating test result	Detection of adenomatous polyps and sessile serrated polyps	Time of day
Ginestra (2023)	Retrospective cohort study	NR	Hospital	USA	HCPs: NR Patients: 1,672	4 – Drug related	Initiation of antimicrobial treatment	Time of day
Hare (2021)	Retrospective cohort study	Primary care clinicians	Primary care	USA	HCPs: NR Patients: 10,757	4 – Drug related	Statin prescribing	Time of day
Harewood (2009)	Prospective study	Gastroenterologists, Surgeons	Hospital, Gastroenterology	Ireland	HCPs: NR Patients: 400	5 – Therapeutic procedure related 2 – Evaluating test result	Cecal intubation (successful completion of colonoscopy) Lesion (polyp, mass) detection	Order number of decision
Hsiang (2019)	Retrospective, observational, quality improvement study	Physicians	Primary care	USA	HCPs: NR Patients: 19,254 (breast cancer screening); 33,468 colorectal cancer screening	1 – Gathering additional information	Ordering of breast cancer screening Ordering of colorectal cancer screening	Time of day

Hueston (1996)	Retrospective study	Physicians, Nurses, Midwives, Obstetricians, Family physicians and Nurse midwives	Hospital	USA	HCPs: NR Patients: 6,440	5 – Therapeutic procedure related	Delivery via caesarean section for foetal distress (vs vaginal delivery)	Time of day
Hughes (2020)	Multi-year analysis (retrospective study)	Physicians	Primary care	USA	HCPs: NR Patients: 77,285	4 – Drug related	Opioid prescribing	Time of day
Hunt (2021)	Retrospective cohort study	Physicians, Urologists, Advanced practice clinicians	Hospital	USA	HCPs: NR Patients: 1,581,826	1 – Gathering additional information	Prostate-specific antigen test ordering	Time of day
Kaneshiro (2010)	Nonrandomized before-and-after study	Gastroenterologists, Gastroenterology trainees	Hospital, Gastroenterology	USA	HCPs: 33 Patients: 778	2 – Evaluating test result	Polyp detection via colonoscopy Adenoma detection via colonoscopy	Time of day
Keswani (2016)	Retrospective patient-level analysis	Gastroenterologists	Hospital, Gastroenterology	USA	HCPs: 18 Patients: 7,004	2 – Evaluating test result	Adenoma detection via colonoscopy	Time of day
Kim (2018)	Retrospective quality improvement study	Physicians, medical assistants	Primary Care	USA	HCPs: NR Patients: 96,291	5 – Therapeutic procedure related	Influenza vaccination	Time of day
Kolla (2023)	Secondary post-hoc analysis of a randomized clinical trial	Physicians	NR	USA	HCPs: 75 Patients: 55,367 encounters with 17,696 patients	9 – Treatment goal	SIC (serious illness communication) conversation (yes/no)	Time into shift
Kushnir (2020)	Retrospective chart review	Nurses	Hospital, Neonatal intensive care unit	USA	HCPs: NR Patients: 101	3 – Defining problem	Neonatal abstinence syndrome scores, resulting from use of the Finnegan scoring tool	Time of day, Time into shift
Lee (2017)	Retrospective cross-sectional study	Gastroenterologists	Ambulatory or inpatient endoscopy centres	USA	HCPs: 126 Patients: 76,445	2 – Evaluating test result	Adenoma detection via colonoscopy	Number of decisions made previously
Leffler (2012)	Prospective cohort study	Gastroenterologists	Hospital, Gastroenterology	USA	HCPs: 32 Patients: 2,193	2 – Evaluating test result	Polyp detection via colonoscopy Adenoma detection via colonoscopy	Time of day, Number of decisions made previously
Linder (2014)	NR – Retrospective study	Physicians	Primary care	USA	HCPs: 204 Patients: 21,867	4 – Drug related	Antibiotic prescribing	Time of day
Long (2011)	Retrospective cross-sectional study	Gastroenterologists	Hospital, Gastrointestinal endoscopy unit	USA	HCPs: NR Patients: 3,421	2 – Evaluating test result	Polyp detection via colonoscopy Adenoma detection via colonoscopy	Time of day, Time into shift
Lu (2023)	Secondary analysis of 2 RCTs	Physicians, Endoscopists	NR	China	HCPs: 6–9 Patients: 1,780	2 – Evaluating test result	Adenoma detection via colonoscopy Polyp detection via colonoscopy	Time of day

(*Continued*)

Table 1. Continued.

Quantitative studies								
Study (First author & Year)	Study Design	Sample (Type of HCPs)	Setting	Country	N (HCP & Patients/ cases)	Decision Type (DICTUM) Not applicable for attributes linked to the decision-making process	Specific Decision(s) *or* Attribute(s) linked to the decision-making process	DF measure(s)
Lurix (2012)	Retrospective chart review	Gastroenterologists	Hospital, Gastroenterology	USA	HCPs: 9 Patients: 3,085	2 – Evaluating test result	Adenoma detection via colonoscopy	Time of day, Order number of decision
Marcondes (2018)	NR – Retrospective study	Gastroenterologists	Private Endoscopy Centre	USA	HCPs: 131 Patients: 86,624	2 – Evaluating test result	Adenoma detection via colonoscopy	Order number of decision
McClelland (2013)	NR – Scenario questionnaire study	Nurses	Hospital, critical care/ cardiac, medical/ surgical, emergency room	USA	HCPs: 65 Patients: 1 hypothetical patient with 20 scenarios	1 – Gathering additional information Not applicable*	Nurse's decision about how likely they are to call a physician based on different scenarios a patient is in Alertness* Stress* Sleepiness*	Before vs after a 12-hour shift (7am – 7pm)
Mitchell (2006)	NR – Retrospective case-control study	Cytotechnologists	Cytology Service; Laboratory and education service specialising in gynaecologic health	Australia	HCPs: NR Patients: 198 false negative reports; 750,805 control screenings	2 – Evaluating test result	Cervical cytology test interpretation decision (proportion of false negative slides)	Time of day
Munson (2011)	Retrospective cohort study	Physicians, Endoscopists	Outpatient endoscopy unit	USA	HCPs: 54 Patients: 3,846	2 – Evaluating test result	Polyp detection via colonoscopy	Time into shift, Time of day
Neprash (2019)	Observational cross-sectional study	Physicians	Primary care	USA	HCPs: 5,603 Patients: 642,262	4 – Drug related 7 – Contact related	Opioid prescribing Nonopioid pain management – prescribing of NSAIDs Referral to a physical therapist Antihypertensive prescribing Statin prescribing	Time of day, Order number of decision
Oakes (2021)	NR – Retrospective study	Cardiologists	Cardiology practice sites	USA	HCPs: NR Patients: 7,271	4 – Drug related	Statin prescribing	Time of day
Paeck (2013)	Retrospective study	Gastroenterologists	Hospital, Gastroenterology	South Korea	HCPs: 63 Patients: 1,293	2 – Evaluating test result	Adenoma detection via colonoscopy Polyp detection via colonoscopy	Time of day, Time into shift
Persson (2019)	Natural experiment	Orthopaedic surgeons	Hospital, Surgery	Sweden	HCPs: 8 Patients: 848	5 – Therapeutic procedure related	Scheduling patient for operation	Order number of decision
Philpot (2018)	NR – Retrospective study	Physicians, Advanced practice providers, Medical residents/ fellows	Primary care	USA	HCPs: NR Patients: 2,772	4 – Drug related	Opioid prescribing	Time of day

Pignatiello (2022)	Secondary analysis of a cross-sectional, descriptive study	Nurses	Sample working in different settings, mainly at academic medical centre or community teaching hospital	USA	HCPs: 160 Patients: NA	Not applicable*	Not applicable – *direct measure of decision fatigue**	DFS Scale
Pu (2019)	Retrospective study	Physicians, Surgeons	Hospital, Gastroenterology	Australia	HCPs: 15 Patients: 558	2 – Evaluating test result	Adenoma detection via colonoscopy Polyp detection via colonoscopy	Time of day
Rawshani (2017)	Cohort study	Emergency Medical Dispatchers	Emergency dispatch centre	Sweden	HCPs: NR Patients: 2,205	7 – Contact related	Assigning case priority	Time of day
Sapci (2022)	NR – Retrospective study	Surgeons	Tertiary referral centre	USA	HCPs: 28 Patients: 13,809	2 – Evaluating test result	Adenoma detection rate (via colonoscopy)	Time of day
Saposnik (2017)	Pilot, double-blinded, parallel-group, randomized clinical trial	Neurologists	Experimental setting; face-to-face meeting of neurologists	Spain	HCPs: 25 Patients: 20 simulated patient-scenarios per HCP	4 – Drug related	Therapeutic decisions for hypothetical MS patients (continue the same treatment, change to a treatment that would not affect the clinical course, or escalating to a more effective agent)	Decisions made in block 2 (10 case-scenarios, after intervention) vs block 1 (10 baselines case-scenarios, before intervention)
Soukup (2019a)	Longitudinal multiphase intervention study	Multidisciplinary oncology teams: Radiologists, Surgeons, Oncologists, Histopathologists, Nurses	Hospital, Oncology	UK	HCPs: 15 Patients: 1,335	3 – Defining problem	Multidisciplinary team decision-making: Quality of presented patient information Quality of disciplinary contribution	Order number of decision
Soukup (2020a)	Multicentre cross-sectional observational study	Multidisciplinary oncology teams: Radiologists, Surgeons, Oncologists, Histopathologists, Nurses	Hospital, Oncology	UK	HCPs: 44 Patients: 822	3 – Defining problem	Multidisciplinary team decision-making: Quality of presented patient information Quality of disciplinary contribution	Order number of decision
Soukup (2020b)	Prospective cross-sectional observational study	Multidisciplinary oncology teams: Radiologists, Surgeons, Oncologists, Histopathologists, Nurses	Hospital, Oncology	UK	HCPs: 44 Patients: 822	3 – Defining problem	Team interaction and communication when deciding on treatment recommendations for patients	Order number of decision
Stecker (2015)	Retrospective Chart Review	Any hospital staff	Hospital	USA	HCPs: NR Patients: 883 stroke alerts	5 – Therapeutic procedure related 4 – Drug related 3 – Defining problem	Activation of hospital stroke alert system Administering tPA (tissue plasminogen activator) as treatment Diagnosing stroke	Time of day
Taylor-Phillips (2015)	Retrospective Review, Re-analysis of existing data	Radiologists, residents, fellows, radiography advanced practitioners	Radiology	UK, USA	HCPs: 118 Patients: 472	2 – Evaluating test result	Abnormality detection in radiography imaging	Order number of decision
Taylor-Phillips (2016)	Randomized Clinical Trial	Radiologists, Radiography advanced practitioners	Breast screening centres	UK	HCPs: 360 Patients: 1.2 million	2 – Evaluating test result	Cancer detection rate (based on reading of digital mammograms) Recall/reader disagreement	Order number of decision

(*Continued*)

Table 1. Continued.

Study (First author & Year)	Study Design	Sample (Type of HCPs)	Setting	Country	N (HCP & Patients/ cases)	Decision Type (DICTUM) Not applicable for attributes linked to the decision-making process	Specific Decision(s) *or* Attribute(s) linked to the decision-making process	DF measure(s)
				Quantitative studies				
Thurtle (2014)	Retrospective study	Gastroenterologists, Colorectal surgeons	Hospital, Gastroenterology	UK	HCPs: 30 Patients: 2,576	2 – Evaluating test result	Polyp detection via colonoscopy	Order number of decision
Trinh (2021)	A retrospective study	Physicians	Outpatient speciality clinics (cardiology, endocrinology, haematology, nephrology, general internal medicine, and rheumatology practices clinics)	USA	HCPs: 20 Patients: 3,342	1 – Gathering additional information 3 – Defining problem	Diagnostic test ordering Diagnostic assessment	First vs last hour of clinic session
Vosshenrich (2021)	Single-centre retrospective study	Radiologists, residents	Hospital, Radiology	Switzerland	HCPs: NR Patients: 117,402	11 – Documentation related	Written radiology report accuracy	Time of day
Wright (2006)	Retrospective study	Certified registered nurse anaesthetist (CRNA), Anaesthesia resident, Staff anaesthesiologist	Hospital, Anaesthesiology	USA	HCPs: NR Patients: 90,159	5 – Therapeutic procedure related	Decisions that have resulted in anaesthetic adverse event (including error events, harm events, other adverse events)	Time of day
Xu (2018)	Retrospective study	Gastroenterologists	Hospital, Gastroenterology	China	HCPs: 19 Patients: 1,342	2 – Evaluating test result	Adenoma detection via colonoscopy	Time of day
Zhang (2023a)	Retrospective study	Endoscopists	Hospital, Endoscopy centre	China	HCPs: 30 Patients: 14,597	2 – Evaluating test result	Lesion detection via esophagogastroduodenoscopy Endoscopy biopsy specimen obtained via esophagogastroduodenoscopy	Time of day
Zhang (2023b)	Retrospective study	Endoscopists	Hospital, Endoscopy centre	China	HCPs: 30 Patients: 12,116	2 – Evaluating test result	Polyp detection via colonoscopy	Time of day
Zheng (2020)	Database analysis	Emergency physicians	Hospital, Emergency	Canada	HCPs: NR Patients: 87,752	7 – Contact related 1 – Gathering additional information	Disposition decisions (whether a patient is discharged or admitted) CT scan ordering (head, chest, abdomen)	Time into shift

Qualitative study

Study (First author & Year)	Methods	Sample (Type of HCPs)	Setting	Country	N (HCPs)	Context	DF relevance
Cygler (2021)	Constructivist grounded theory	Internal medicine residents	Hospital, Clinical teaching unit	Canada	17	Perceived benefits and drawbacks of 24-hour in-house call	Decision fatigue emerged as a theme rather than specifically being addressed by the study question

Opinions, editorials, book chapters, or narrative overviews

Study (First author & Year)	Article Type	Population of interest (Type of HCPs)	Setting	Country	DF relevance
Dubash (2020)	Editorial	Emergency physicians	Hospital, Emergency	Australia	Decision fatigue is the focus of the article. Individual and departmental factors contributing to DF are discussed. Manifestations of decision fatigue are described, and potential solution are suggested.
Hatami (2022)	Editorial	Nurses	NR	No geographical focus	Decision fatigue is the focus of the article. It is discussed in the context of the COVID-19 pandemic which put additional strains on nurses, both on terms of physical and mental workload. The article lacks a clear aim and makes only generic statements relating to decision fatigue in a pandemic situation.
Lee (2013)	Narrative Overview	Radiologists, Radiology residents	Hospital, Radiology	USA	Decision fatigue is considered amongst other factors in the broader context of cognitive and system factors that contribute to diagnostic errors in radiology.
Lee (2021)	Book chapter	Clinicians	Hospital, Emergency	No geographical focus	Decision fatigue is the focus of the book chapter. Four case studies are presented that describe situations where clinicians become decision fatigued and the effects of that are described: poor communication patterns, choosing the easy option and avoiding difficult conversation, oversimplification and anchoring to what seemed the obvious diagnosis, ordering unnecessary diagnostic imaging. A list of potential solutions is provided.
Moorhouse (2020)	Opinion Piece	Physicians	Primary care	UK	Decision fatigue is the focus of the article. It is discussed in relation to burnout and dedicated breaks are suggested as a strategy for limiting DF. A list of tips to ameliorate DF is also provided.
Nasa (2023)	Editorial	Clinicians	Hospital, Emergency	No geographical focus	Decision fatigue is the focus of the article. Associated factors such as working hours, workload, work schedule, lack of resources, multitasking, and burnout are discussed. Effects of decision fatigue on patient care but also personal life are listed. Finally mitigating strategies are discussed.
Oakes (2021)	Opinion Piece	Clinicians generally	Health care generally	USA	Disparities in care by appointment time are discussed with decision fatigue given as the main explanation. Different potential ways for potentially limiting negative impacts of DF on care are proposed.
Schweitzer (2023)	Opinion Piece	Clinicians, Physicians, Nurses, Physiotherapists	Hospital, Emergency and Neurosurgical Intensive Care Unit (ICU)	Australia	Decision fatigue and limited willpower is the focus of the article. A conceptual framework is introduced to visualise the connection between internal resources, external resources, willpower, self-control and clinical decision fatigue. Three case studies are described to illustrate that depleted willpower can affect and compromise healthcare delivery. Research areas for addressing decision fatigue in the healthcare context are proposed.

Systematic review

Study (First author & Year)	Article Type	Population of interest (Type of HCPs)	Setting	Country	DF relevance
Soukup (2019b)	Integrated Literature Review	Multidisciplinary oncology teams: Radiologists, Surgeons, Oncologists, Histopathologists, Nurses	Hospital, Oncology	Worldwide review, but restricted to English language publications	The term 'decision fatigue' is not used as such, however the review's focus is on how time-on-task affects decision-making. It is explored how and why time spent on a task affects performance, and what strategies can be employed by cancer teams to counteract negative effects and improve quality and safety.

Of the 72 quantitative studies, three were randomised controlled trials (Bersani et al., 2020; Saposnik et al., 2017; Taylor-Phillips et al., 2016), the rest were non-randomised studies (Addis et al., 2020; Agarwal et al., 2022; Allan et al., 2019; Almadi et al., 2015; Anastasian et al., 2014; Backmann et al., 2021b, 2021a; Beg et al., 2021; Bernstein et al., 2022; Bonafide et al., 2017; Canfield et al., 2020; Castrén et al., 2015; Chan et al., 2009; Changolkar et al., 2020; Cheng et al., 2011; Christensen et al., 1977; Dai et al., 2015; Delaney & Fogg, 2005; Dexheimer et al., 2017; Dominello et al., 2015; Edwards et al., 2003; Ergün-Şahin et al., 2022; Evans et al., 1984; Faulk et al., 2013; Freedman et al., 2011; Ginestra et al., 2023; Hare et al., 2021; Harewood et al., 2009; Hsiang et al., 2019; Hueston et al., 1996; Hughes et al., 2020; Hunt et al., 2021; Kaneshiro et al., 2010; Keswani et al., 2016; Kim et al., 2018; Kolla et al., 2023; Kushnir et al., 2020; A. Lee et al., 2017; Leffler et al., 2012; Linder et al., 2014; Long et al., 2011; Lu et al., 2023; Lurix et al., 2012; Marcondes et al., 2018; McClelland et al., 2013; Mitchell et al., 2006; Munson et al., 2011; Neprash & Barnett, 2019; Oakes et al., 2021; Paeck et al., 2013; Persson et al., 2019; Philpot et al., 2018; Pignatiello et al., 2022; Pu et al., 2019; Rawshani et al., 2017; Sapci et al., 2022; Soukup, Gandamihardja, et al., 2019; Soukup, Lamb, Morbi, et al., 2020; Soukup, Lamb, Shah, et al., 2020b; Stecker et al., 2015; Taylor-Phillips et al., 2015; Thurtle et al., 2014; Trinh et al., 2021; Vosshenrich et al., 2021; Wright et al., 2006; Xu et al., 2018; Zhang, Chen, Cao, et al., 2023; Zhang, Chen, Wang, et al., 2023b; Zheng et al., 2020). Nearly two-thirds (45/72) of the quantitative studies were conducted in the USA, 11% (8/72) in the UK, 7% in China (5/72), and 4% (3/72) in both Sweden and Australia. Other geographical locations were Canada ($n = 2$), Ireland ($n = 1$), Norway ($n = 2$), South Korea ($n = 1$), Spain ($n = 1$), Switzerland ($n = 1$), and Turkey ($n = 1$). Nearly two-thirds ($n = 45$) of the studies were conducted in a hospital setting, eight in a primary care setting, and further settings included but were not limited to a medical telephone helpline, endoscopy/ gastroenterology centres, and emergency medical communication centres.

Forty-five of 72 studies reported the number of healthcare professionals (HCPs)/decision makers involved, and of those that did, numbers ranged from 3 to 5,603 HCPs. The number of patients involved in studies ranged from 67 to over 21 million. However, most studies did not clarify whether the reported numbers of patients referred to individual patients or individual decisions. For example, it was often unclear whether the numbers reported included revisits or included multiple decisions per patient by each HCP (for example, in mammography reading decisions where, generally, two images are included per patient). We still refer to 'patients' in this report for ease.

The only qualitative study included in this review (Cygler et al., 2021) was conducted in Canada, and had 17 study participants who were all internal medicine residents in clinical teaching units at hospitals.

Three opinion pieces (Moorhouse, 2020; Oakes & Patel, 2021; Schweitzer et al., 2023), three editorials (Dubash et al., 2020; Hatami et al., 2022; Nasa & Majeed, 2023), one book chapter (X. Q. Lee, 2021) and one narrative overview (C. S. Lee et al., 2013) were included: two from the USA, two from Australia, one from the UK, and three with no geographical focus. The populations of interest were emergency physicians, nurses, physiotherapists, radiologists, primary care physicians, and clinicians in general.

The included literature review (Soukup, Lamb, et al., 2019b) had no geographic inclusion restrictions, but only considered English-language publications. Its population and setting of interest were specifically multidisciplinary oncology teams in hospitals.

Decision fatigue effects were investigated quantitatively in more than 57 different types of medical decisions (or observable decision outcomes). When categorised using the revised DICTUM, the majority of decisions studied fell into DICTUM categories DC2 (Evaluating test results), DC4 (Drug related), and DC5 (Therapeutic procedure related).

DC1 decisions (Gathering additional information) consisted of decisions to order screening (Hsiang et al., 2019) or diagnostic tests (Ergün-Şahin et al., 2022; Hunt et al., 2021; Trinh et al., 2021; Zheng et al., 2020), the decision to call another HCP to consult (McClelland et al., 2013), and the use of a patient safety dashboard to look at additional information (Bersani et al., 2020). DC2 decisions (Evaluating test results) contained evaluative decisions made during endoscopy procedures (Almadi et al., 2015; Beg et al., 2021; Chan et al., 2009; Freedman et al., 2011; Harewood et al., 2009; Kaneshiro et al., 2010; Keswani et al., 2016; A. Lee et al., 2017; Leffler et al., 2012; Long et al., 2011; Lu et al., 2023; Lurix et al., 2012; Marcondes et al., 2018; Munson et al., 2011; Paeck et al., 2013; Pu et al., 2019; Saposnik et al., 2017; Thurtle et al., 2014; Xu et al., 2018; Zhang, Chen, Cao, et al., 2023; Zhang, Chen, Wang, et al., 2023b), mammography (Backmann et al., 2021b, 2021a; Bernstein et al., 2022; Taylor-Phillips et al., 2016) and radiograph reading decisions (Christensen et al., 1977), cervical cytology test interpretation decisions (Mitchell et al., 2006), and responding to physiologic monitor alarms (Bonafide et al., 2017). DC3 decisions (Defining problem) comprised temperature recording (Edwards et al., 2003), using a neonatal condition scoring tool (Kushnir et al., 2020), diagnosing strokes (Stecker et al., 2015), multidisciplinary oncology team decision-making (Soukup, Gandamihardja, et al., 2019; Soukup, Lamb, Morbi, et al., 2020; Soukup, Lamb, Shah, et al., 2020b), and making diagnostic assessments (Trinh et al., 2021). DC4 Drug related decisions covered prescribing decisions of drugs such as antibiotics (Linder et al., 2014), antihypertensives (Neprash & Barnett, 2019), nonopioid pain management medications (Neprash & Barnett, 2019), opioids (Agarwal et al., 2022; Hughes et al., 2020; Neprash & Barnett, 2019; Philpot et al., 2018), and statins (Hare et al., 2021; Neprash & Barnett, 2019; Oakes et al., 2021). They also included initiation of antimicrobial treatment (Ginestra et al., 2023), administration of antipyretics (Edwards et al., 2003) and tissue plasminogen activator treatment (Stecker et al., 2015), medication order alert responding (Dexheimer et al., 2017), and therapeutic medication decisions for hypothetical MS patients (Saposnik et al., 2017). DC5 Therapeutic related decisions included the transfusion of different blood products during cardiac surgery (Addis et al., 2020), delaying extubation at the end of multi-level prone spine surgery (Anastasian et al., 2014), administering influenza vaccinations (Changolkar et al., 2020; Kim et al., 2018), restraining a patient (Delaney & Fogg, 2005), delivering babies via caesarean section (Evans et al., 1984; Hueston et al., 1996), successful completion of colonoscopy (cecal intubation) (Harewood et al., 2009), scheduling patients for surgery (Persson et al., 2019), activating the hospital stroke alert system (Stecker et al., 2015), and decisions that resulted in radiation treatment errors (Dominello et al., 2015) or anaesthetic adverse events (Wright et al., 2006). No decisions that would fall into DC6 (Legal and insurance related) were identified. DC7 Contact related decisions included decisions to refer patients to a physical therapist (Neprash & Barnett, 2019), admitting patients to an inpatient rehabilitation facility (Faulk et al., 2013), disposition decisions (decisions about whether the patient is discharged or admitted) (Zheng et al., 2020), pre-hospital assessment decisions (Castrén et al., 2015), assigning case priority (Rawshani et al., 2017), and triage decisions via medical telephone helpline (Allan et al., 2019). No DC8 Advice and precaution decisions were identified. A single DC9 Treatment goal decision was identified; the decision of whether to conduct a serious illness communication conversation (Kolla et al., 2023). No DC10 deferment decisions were identified. DC11 Documentation related decisions included decisions to enter a patient order into the electronic health record (Canfield et al., 2020), and radiology report accuracy (Vosshenrich et al., 2021). Finally, DC12 Infection prevention decisions were related to hand hygiene compliance decisions (Cheng et al., 2011; Dai et al., 2015).

Further attributes investigated that are linked to medical decision-making processes but which are not decisions and so could not be categorised into the revised DICTUM included a fatigue score, psychomotor vigilance (Beg et al., 2021), alertness, stress, sleepiness (McClelland et al., 2013), and also DF-scale measurements that were linked to providing competent nursing practice (Pignatiello et al., 2022).

Operationalisation and measurement of decision fatigue

Decision fatigue was operationalised and measured in various ways by the studies included, and some studies used multiple ways of measurement. The majority of studies used a time-related proxy measure for decision fatigue: 47 studies analyse decision-making in relation to the time of day (Addis et al., 2020; Anastasian et al., 2014; Backmann et al., 2021a; Bernstein et al., 2022; Bersani et al., 2020; Canfield et al., 2020; Castrén et al., 2015; Chan et al., 2009; Changolkar et al., 2020; Cheng et al., 2011; Delaney & Fogg, 2005; Dexheimer et al., 2017; Dominello et al., 2015; Edwards et al., 2003; Evans et al., 1984; Faulk et al., 2013; Freedman et al., 2011; Ginestra et al., 2023; Hare et al., 2021; Hsiang et al., 2019; Hueston et al., 1996; Hughes et al., 2020; Hunt et al., 2021; Kaneshiro et al., 2010; Keswani et al., 2016; Kim et al., 2018; Kushnir et al., 2020; Kushnir et al., 2020; Leffler et al., 2012; Linder et al., 2014; Long et al., 2011; Lu et al., 2023; Lurix et al., 2012; Mitchell et al., 2006; Munson et al., 2011; Neprash & Barnett, 2019; Oakes et al., 2021; Paeck et al., 2013; Philpot et al., 2018; Pu et al., 2019; Rawshani et al., 2017; Sapci et al., 2022; Stecker et al., 2015; Vosshenrich et al., 2021; Wright et al., 2006; Xu et al., 2018; Zhang, Chen, Cao, et al., 2023; Zhang, Chen, Wang, et al., 2023b), 12 studies use time into shift (Agarwal et al., 2022; Allan et al., 2019; Almadi et al., 2015; Bonafide et al., 2017; Christensen et al., 1977; Dai et al., 2015; Kolla et al., 2023; Kushnir et al., 2020; Long et al., 2011; Munson et al., 2011; Paeck et al., 2013; Zheng et al., 2020), one study used time since the last break (Allan et al., 2019), one study used hours off since the previous shift (Dai et al., 2015), and three studies compared decisions made at earlier blocks of time with later blocks of time (McClelland et al., 2013; Saposnik et al., 2017; Trinh et al., 2021). Two studies integrated the time of day measure with time into the shift (Long et al., 2011; Munson et al., 2011). Eighteen studies used an ordinal decision type variable, analysing decision-making according to either order number of decision (Backmann et al., 2021a; Harewood et al., 2009; Lurix et al., 2012; Marcondes et al., 2018; Neprash & Barnett, 2019; Persson et al., 2019; Soukup, Gandamihardja, et al., 2019; Soukup, Lamb, Morbi, et al., 2020; Soukup, Lamb, Shah, et al., 2020b; Taylor-Phillips et al., 2015, 2016; Thurtle et al., 2014) or the number of decisions made previously (Allan et al., 2019; Almadi et al., 2015; Ergün-Şahin et al., 2022; A. Lee et al., 2017; Leffler et al., 2012). One study utilised the decisional fatigue scale (Pignatiello et al., 2022).

Decision fatigue effects

Table 2 presents the summary details of studies, organised by DICTUM categories, decision type, decision fatigue measure, and the measured effects of decision fatigue on the decision/outcome. It also indicates whether study results provided support for the decision fatigue hypothesis. As some studies investigated several different types of medical decisions and/or used multiple decision fatigue measures, the number of studies does not correspond with the number of cases of formal tests of decision fatigue, and so the same study reference can appear in different DICTUM categories and more than once within categories. Across all included quantitative studies, statistically significant support for the decision fatigue hypothesis was found in 58 of 130 (45%) cases, no-support, or statistically non-significant results were found in 42 cases (32%), and inconclusive results (i.e., results that were neither decisively in support nor contradicted the hypothesis) in 30 cases (23%).

DC1: Gathering Additional Information. The decision fatigue hypothesis was examined in seven studies (Bersani et al., 2020; Ergün-Şahin et al., 2022; Hsiang et al., 2019; Hunt et al., 2021; McClelland et al., 2013; Trinh et al., 2021; Zheng et al., 2020) including a total of 12 cases of formal tests of decision fatigue. Eight cases supported the hypothesis (Ergün-Şahin et al., 2022; Hsiang et al., 2019; Hunt et al., 2021; Trinh et al., 2021; Zheng et al., 2020), two remained inconclusive (Bersani et al., 2020; McClelland et al., 2013), and two did not support the hypothesis (Ergün-Şahin et al., 2022; Zheng et al., 2020).

The studies that showed significant support demonstrated that the number of diagnostic tests ordered per patient and the probability of ordering an imaging test decreased for each additional patient seen (Ergün-Şahin et al., 2022), and that ordering rates of breast screening (Hsiang et al., 2019), colorectal screening (Hsiang et al., 2019), CT head (Zheng et al., 2020) and abdomen scans (Zheng et al., 2020) and prostate antigen tests (Hunt et al., 2021) declined significantly over a shift/day.

Results were inconclusive for nurses' decisions to call a physician in response to specific patient symptoms before versus after a shift (McClelland et al., 2013), and for decisions to use a patient safety dashboard (Bersani et al., 2020). For both, significant changes across the shift or the day were observed, however, the nature of those changes was not clearly (contra-) indicative of decision fatigue.

Table 2. Results of included quantitative studies (ordered by DICTUM category).

Decisions/Decision Outcomes	DF Measure(s)	Measured effects of DF on Decision/Outcome	Support for DF hypothesis
1 – Gathering additional information			**8** / **2** / **2**
Diagnostic test ordering	Number of decisions made previously	For each additional examined patient, physicians decrease the number of diagnostic test orders per patient.[23]	Yes
		For each additional examined patient, the probability of a diagnostic test order does not change significantly.[23]	No
		For each additional examined patient, the probability of a lab test order does not change significantly.[23]	No
		For each additional examined patient, the probability of an imaging test order does decrease. [23]	Yes
	First vs last hour of clinic session	Sixteen physicians had no statistically significant differences between the first and last hour in numbers of tests ordered. Four physicians had statistically significant differences in diagnostic tests ordered per patient encounter between the first and last hours of their day.[66]	Yes, partly
Ordering of breast cancer screening	Time of day	Relative to 8 AM, the adjusted odds ratio (OR) of clinician ordering of breast cancer screening was significantly lower for each hour from 10 AM to 5 PM.[30]	Yes
Ordering of colorectal cancer screening	Time of day	Relative to 8 AM, the adjusted OR of clinician ordering of colorectal cancer screening was significantly lower for each hour from 10 AM to 5 PM.[30]	Yes
Prostate-specific antigen test ordering	Time of day	Overall decline in PSA testing rates as outpatient appointments progressed throughout the day.[33] Decline was present in both appropriate and low-value settings, but proportionately greater impact was observed when testing was appropriate.[33]	Yes
Nurses' decision about how likely they are to call a physician	Before vs after a 12-hour shift (7am–7pm)	Nurses showed significant changes across the shift in their willingness to call a physician given a particular patient's information or cues.[46]	Unclear
CT scan ordering (head, chest, abdomen)	Time into shift	There was a significant decline in the hourly rates of CT head ordering as the shift went on.[72]	Yes
		CT chest ordering did not change significantly over the shift.[72]	No
		There was a significant decline in the hourly rates of CT abdomen ordering as the shift went on.[72]	Yes
Use of a Patient Safety Dashboard	Time of day	High concentration of logins between 5 and 8 am, large increase during morning rounds (8:30–11:30 am) and decrease each subsequent hour.[10]	Unclear
2 – Evaluating test result			**21** / **19** / **11**
Endoscopy reading decisions			
Adenoma detection	Time into shift	Adenoma detection rate decreased with an increase in number of scoping hours elapsed until index colonoscopy.[4] Statistically significant difference in ADR when comparing colonoscopies performed ≤3 h versus >3 h after the start of the endoscopy session.[4]	Yes
		Colonoscopies performed during the 4th hour of a session were associated with a decreased ADR.[51]	Yes
	Number of decisions made previously	The 'no adenoma' group had a higher number of colonoscopies and a higher number of endoscopic procedures performed before the index colonoscopy.[4]	Yes
		No inverse relationship between the raw, consensus, and RVU fatigue scores and the frequency of adenoma detection.[39]	No
		No significant variation in adenoma detection rate, or advanced adenoma detection rate.[40]	No
		In post hoc analysis, adenoma detection rates appeared lower after the fifth case of the day for endoscopists with low volumes of cases and after the tenth case of the day for endoscopists with high volumes of cases.[40]	Yes
	Time of day	Significant negative relationship between start time and adenoma yield.[34]	Yes
		Adenoma detection was not statistically significantly associated with time of day.[35]	No
		No significant variation in adenoma detection rate, or advanced adenoma detection rate.[40]	No

(*Continued*)

Table 2. Continued.

Decisions/Decision Outcomes	DF Measure(s)	Measured effects of DF on Decision/Outcome	Support for DF hypothesis
		No decrease in ADR was noted with an increasing time of day.[44]	No
		There was a trend toward a declining ADR for each subsequent hour of the day, although there was no statistical significance.[51]	No
		There is a statistically significant association between ADR and time as a continuous variable.[55]	Yes
		Adenoma detection rate does not decrease with the time of day. ADR increases for male patients from roughly 29% at 7am to 43% at 5pm. For female patients the ADR fluctuates less.[61]	No
		Adenoma detection rate upraised over the time. The first climax for ADR was presented at 09:00–10:00. After the lunch break, ADR rose again and persisted to the end of the day. A significant inclined trend in ADR was noted for each hour blocks of a full day.[69]	Yes
		Adenoma detection rate declines significantly over a half-day of work.[43]	Yes
	Time of day & time into shift	Progression of time was significantly associated with reduced odds of adenoma detection.[42] Reductions with regards to shift timing or duration did not reach statistical significance.[42]	Unclear
	Order number of decision	No decrease in ADR was noted with an increasing queue position.[44]	No
		ADR decreases after the 7th colonoscopy within a procedure day.[45]	Yes
Lesion/polyp detection	Time of day	Polyp yield diminished as day progressed, with no difference in proportion of hyperplastic versus adenomatous polyps.[14]	Yes
		Polyp detection was independent of time of day.[26]	No
		Start time significantly predicted decreased polyp yield.[34]	Yes
		No significant variation in polyp detection rate.[40]	No
		There is no clear trend in the way polyp detection rate changes over time.[43]	Unclear
		There was a trend toward a declining PDR for each subsequent hour of the day, although there was no statistical significance.[51]	No
		Lesion detection rate decreases significantly as the day progresses.[70]	Yes
		Endoscopy biopsy rate decreases significantly as the day progresses.[70]	Yes
		Polyp detection rate and polyps detected per colonoscopy decrease significantly as the day progresses.[71]	Yes
	Time of day & time into shift	For each advancing hour of the day, there was a significant reduction in odds of polyp detection.[42] This same reduction in the odds of polyp detection was seen regardless of shift timing or duration.[42]	Yes
		Polyp detection rate (PDR) does not decline steadily during the day.[48] PDR declines only during the afternoon shift (1:30–4:30 pm) but remains stable otherwise.[48]	Unclear
	Order number of decision	No significant differences in correct lesion detection in association with reading order.[8]	No
		Detection rates remained stable on A.M. and P.M. lists.[29]	No
		Polyp detection rate was not shown to be affected by rank order of colonoscopy within a list.[65]	No
	Number of decisions made previously	No significant variation in polyp detection rate.[40] In post hoc analysis, polyp detection rates appeared lower after the fifth case of the day for endoscopists with low volumes of cases and after the tenth case of the day for endoscopists with high volumes of cases.[40]	Unclear
Endoscopy reading time	Order number of decision	A reduction in reading time was observed; reading times were on average 25% faster on reading Case 6 compared to Case 1.[8]	Yes
Mammography reading decision	Time of day	Percentage of positive readings was highest before lunch and in the early afternoon (4.9%): False positive was highest in both periods (4.5%) and true positive was highest in the early afternoon (0.5%).[6]	Unclear

(*Continued*)

Table 2. Continued.

Decisions/Decision Outcomes	DF Measure(s)	Measured effects of DF on Decision/Outcome	Support for DF hypothesis
		Percentage of true negative was highest in the evening (95.6%), and of false negative was highest at lunchtime (0.07%).[6]	
		Gradually decreasing predicted sensitivity (true positive rate) throughout the day.[6]	Yes
		Gradually increasing predicted specificity (true negative rate) throughout the day.[6]	No
		The odds of a false positive finding increased for every additional hour of reading time (both for digital breast tomosynthesis (more effortful) and for digital mammography).[9]	Yes
		The odds of a true positive finding increased for every additional hour of reading time (both for digital breast tomosynthesis (more effortful) and for digital mammography).[9]	No
		The odds of a recall increased for every additional hour of reading time (both for digital breast tomosynthesis (more effortful) and for digital mammography).[9]	Yes
	Order number of decision	True and false positives were negatively associated with image position within batch.[7]	Unclear
		Rates of true and false negatives were positively associated with image position within batch (the second was not statistically significant).[7]	Unclear
		Cancer detection rate for individual readers did not change with time spent on task, as represented by near identical odds of detecting cancer between the first and 40th case.[64]	No
		Recall rate for individual readers reduced with time on task.[64]	Yes
Response to physiologic monitor alarms	Time into shift	With each successive hour that passed in a nurse's shift, response time to physiologic monitor alarms was slower.[11]	Yes
Radiograph reading decisions: Detection of pulmonary nodules	Time into shift	5 residents' performance improved in fatigued state, 8 declined, and 1 remained unchanged. Difference was very small in almost every case. Average of 44.8 nodules (61.4%) were detected in rested state, and 44.3 nodules (60.7%) in fatigued state.[17]	Unclear
Cervical cytology test interpretation decision	Time of day	False negative rates may be lowest after a break; before 9 AM and between 1 and 2 PM but did not reach statistical significance.[47]	No
Abnormality detection in radiography imaging	Order number of decision	There was a sensitivity decrease or specificity increase over the course of reading:	
		100 chest x-rays[63]	Unclear
		60 bone fracture x-rays[63]	Unclear
		100 chest CT scans.[63]	Unclear
		This effect was not found in the shorter mammography sessions with 27 or 50 cases.[63]	No
3 – Defining problem			**9 9 2**
Temperature recording	Time of day	A strong second hourly pattern emerged with temperatures recorded more frequently between 19:00 and 05:00 and less frequently between 07:00 and 17:00.[22]	Unclear
Using the Finnegan scoring tool	Time of day, Time into shift	No relationship between the scoring observer's shift hour and daily high Neonatal abstinence syndrome (NAS) scores.[38]	No
		Mean high NAS scores decreased as time of day increased.[38] Daily high NAS score tended to occur mostly between 12:00 and 15:00 PM, which is in the middle of the observer's shift.[38]	Unclear
		When comparing the calculated NAS score using only objective data points to the total score, no significant difference occurred throughout the day.[38]	No
Multidisciplinary team decision-making: Quality of presented patient information	Order number of decision	Serial position of cases in meetings was negatively correlated with information quality.[58]	Yes
		Serial position of treatment decisions in meetings was negatively correlated with information quality.[59]	Yes
Quality of disciplinary contribution	Order number of decision	Serial positions of cases in meetings were negatively correlated with contribution quality.[58]	Yes
		Serial position of treatment decisions in meetings was negatively correlated with contribution quality.[59]	Yes

(*Continued*)

Table 2. Continued.

Decisions/Decision Outcomes	DF Measure(s)	Measured effects of DF on Decision/Outcome	Support for DF hypothesis
Team interaction and communication when deciding on treatment recommendations for patients	Order number of decision	Serial position of treatment decisions in meetings was negatively correlated with asking questions and positive reactions.[59]	Yes
		Serial position of treatment decisions in meetings was not significantly correlated with giving answers and negative reactions.[59]	No
		Serial position of treatment decisions in meetings was negatively correlated with:	Yes
		• showing solidarity and agreeing (positive socio-emotional interactions)[60]	
		• showing tension (negative socio-emotional interactions)[60]	No
		• asking for orientation & opinion (task-oriented communication – asking questions)[60]	Yes
		Serial position of treatment decisions in meetings was positively correlated with:	No
		• Tension release (positive socio-emotional interactions)[60]	
		• Disagreeing (negative socio-emotional interactions)[60]	Yes
		• Giving opinions (task-oriented communication – giving answers)[60]	No
		Serial position of treatment decisions in meetings was not significantly correlated with:	No
		• Showing antagonism (negative socio-emotional interactions)[60]	
		• Giving orientation & suggestions (task-oriented communication – giving answers)[60]	No
Diagnoses assessed	First vs last hour of clinic session	Fourteen physicians had no statistically significant differences between the first and last hour in numbers of diagnoses assessed. Six physicians had statistically significant.[66]	Yes, partly
Diagnosing stroke	Time of day	There was no effect of the time of day on the rate at which the diagnosis of stroke made.[62]	No
4 – Drug related			7 / 3 / 5
Response to medication order alert	Time of day	For every one-hour increase in clock time, salience probability change is observed. Salience probability increases from 07:00 until midnight and then decreases again until 07:00.[20]	Unclear
Antipyretic administration	Time of day	Antipyretics were administered throughout the day with most frequent administrations at 08:00 and 16:00 and a more even distribution between 02:00 and 06:00.[22]	Unclear
Initiation of antimicrobial treatment	Time of day	No clear pattern; decline in probabilities of initiating antimicrobials throughout the night shift.[27]	Unclear
Statin prescribing	Time of day	Compared with 8 AM, statin prescribing was significantly lower at all hours except 9 AM. The overall time trend for statin prescribing decreased significantly.[28]	Yes
		New statin prescriptions significantly declined as the clinic day progressed, with the largest reduction occurring between 12 and 2 PM.[50]	Yes
	Order number of decision	Initial orders for statins did not increase as a function of appointment order.[49]	Unclear
Opioid prescribing	Time of day	Hourly likelihoods of patients being prescribed opioids increased significantly throughout the clinical day.[32]	Yes
		Appointments for lower back pain occurring later in the day had significantly higher odds of resulting in an opioid prescription.[53]	Yes
		Increasing rates of opioid prescribing as appointments progressed through the day.[49]	Yes
	Time into shift	Among visits for which opioids are sometimes indicated, prescribing increased per hour on shift.[2] Among visits for which opioids are usually not indicated, opioid prescriptions fell per hour on shift. [2]	Unclear
Antibiotic prescribing	Time of day	Antibiotic prescribing increased throughout the morning and afternoon clinic sessions for antibiotics sometimes indicated and antibiotics never indicated ARIs.[41]	Yes
Nonopioid pain management – prescribing of NSAIDs	Time of day	No association between appointment timing and NSAID prescribing.[49]	No
Antihypertensive prescribing	Order number of decision	Initial orders for antihypertensives did not increase as a function of appointment order.[49]	No
Therapeutic decisions for hypothetical MS patients (continue the same treatment, change to a treatment that would not affect the clinical course, or escalating to a more effective agent)	Decisions made in block 2 (10 case-scenarios, after intervention) vs block 1 (10 baselines case-scenarios, before intervention)	There was increased prevalence of therapeutic inertia in the second set of case-scenarios among participants in the control group, but not in the active group.[57]	Yes
Administering tPA (tissue plasminogen activator) as treatment	Time of day	There was no effect of the time of day on the rate at which tPA was given.[62]	No

(*Continued*)

Table 2. Continued.

Decisions/Decision Outcomes	DF Measure(s)	Measured effects of DF on Decision/Outcome	Support for DF hypothesis
5 – Therapeutic procedure related			**6** **4** **6**
Blood product transfusion during cardiac surgery	Time of day	Probability of patient receiving a transfusion changed significantly with later case start times.[1]	Unclear
Red blood cell transfusion	Time of day	Probability of receiving an intraoperative red blood cell transfusion increased with later case start times. [1]	Yes
Transfusion for plasma	Time of day	No difference in probability of transfusion for plasma.[1]	No
Transfusion for cryoprecipitate	Time of day	No difference in probability of transfusion for cryoprecipitate.[1]	No
Transfusion for platelets	Time of day	No difference in probability of transfusion for platelets.[1]	No
Delay extubating at the end of multi-level prone spine surgery	Time of day	Case end time is an independent factor that correlates with delayed extubation in multilevel spine surgery.[5]	Yes
Influenza vaccination	Time of day	Appointment time of day was associated with a decrease in vaccination rates.[15]	Yes
		Relative to the 8 AM, adjusted odds ratios of vaccination were significantly lower for each subsequent hour of the day and for overall linear trend.[36]	Yes
Restraining a patient	Time of day	Clustering of restraint incidents was found during particular times of day. Seventy-nine of restraint incidences occurred in the early afternoon (23 per cent), in the mid-afternoon (22 per cent), and in the early evening (22 per cent). Restraint incidents rarely occurred in the early morning or during sleeping hours.[19]	Unclear
Radiation treatment delivery errors	Time of day	For primary treatment delivery errors, there was no significant difference in error rate based on time of day.[21]	No
Delivery via caesarean section vs vaginal delivery	Time of day	Frequencies of vaginal deliveries varied by time of day. While appearing similar from 2:00 AM to 2:00 PM, frequencies gradually fell throughout the evening, reaching a lowest point from 8:00 PM to 10:00 PM[24]	Unclear
		Caesarean section for foetal distress was relatively constant between 3:00 AM and 8:59 PM, but then showed an increase between the hours of 9:00 PM and 2:59 AM.[31]	Unclear
Cecal intubation (successful completion of colonoscopy)	Order number of decision	On A.M. lists, completion rates decline with successive procedures. Completion rates for 1st to 3rd procedures (90%) were significantly higher than for 4th and subsequent procedures (76%) ($P = 0.03$).[29] There was no difference on P.M. lists: 87% (1st), 86% (2nd), and 92% (3rd and subsequent).[29]	Unclear
Scheduling patient for operation	Order number of decision	Patients who met surgeon toward the end of their shift were less likely to be scheduled for an operation compared with those who were seen first.[52]	Yes
Activation of hospital stroke alert system	Time of day	Frequency of stroke alerts varied throughout the day, being most frequent between 9 am and 6 pm There was an increase from 6:00–12:00, a decrease until 16:00, an increase until 18:00 and a decrease until 5:00.[62]	Unclear
Decisions that have resulted in anaesthetic adverse event	Time of day	Compared with a reference start hour of 7 am, adverse events were more frequent for cases starting during the 3 and 4 pm hours. Post hoc inspection of data revealed that the predicted probability increased from a low of 1.0% at 9 am to a high of 4.2% at 4 pm.[68]	Yes
6 – Legal and insurance related			**x** **x** **x**
7 – Contact related			**3** **4** **2**
Triage decisions (referral within 12 hr vs outside 12-hr window)	Number of decisions made previously	For every additional call taken since nurses' last break, the odds of making a conservative decision increased.[3]	Yes
		No effect of number of decisions since the start of the shift.[3]	No
	Time into shift	No effect of time elapsed since start of shift.[3]	No
	Time since break	The odds of making a conservative decision increased for every hour that passed since the last break.[3]	Yes
Pre-hospital assessment decision	Time of day	Frequency of patients with non-specific complaints (NSCs) peaked between approximately 10.00 am and 2.00 pm while the frequency of patients with specific conditions had its peak at approximately 9 am as well as the peak for all EMS transportations starts around 10.00 am[13]	Unclear

(*Continued*)

Table 2. Continued.

Decisions/Decision Outcomes	DF Measure(s)	Measured effects of DF on Decision/Outcome	Support for DF hypothesis		
Admission to inpatient rehabilitation facility	Time of day	Patients admitted later in the day are more likely to RTACH (return to the acute care hospital) than are patients admitted earlier in the day.[25]		Yes	
Referral to a physical therapist	Time of day	Ordering of referral to physical therapy was without a statistically significant increase as the day progressed.[49]		No	
Assigning case priority	Time of day	Probability of missing to give priority 1 increased during the morning and reached a maximum at 12.59 PM. The risk declined thereafter.[56]		Unclear	
Disposition decisions (whether a patient is discharged or admitted)	Time into shift	No clinically important difference in consultations requested over the eight-hour shift. Percentage of patients receiving consultations was 47.0% in the first hour and 47.4% in the last hour.[72] No important difference in patients discharged without consultation over the shift. Percentage of patients discharged without consultation was 52.9% in the first hour and 52.4% in the last hour.[72]		No	
8 – Advice and precaution			**x**	**x**	**x**
9 – Treatment goal			**1**	**0**	**0**
Serious illness communication conversation	Time into shift	Oncology clinicians' likelihood of having and documenting serious illness conversations decreases as a clinic session progresses.[37]	Yes		
10 – Deferment			**x**	**x**	**x**
11 – Documentation related			**1**	**1**	**1**
Patient order entry/retraction into electronic health records	Time of day	No clear pattern of order retraction by hour is observable.[12]		No	
		Time of day appears to be a factor with wrong patient order entry rates being slightly lower during the night (1900–0700) than the day.[12]		Unclear	
Written radiology report accuracy	Time of day	Decreasing report similarity with increasing work hours was observed.[67] Temporary increases in report similarity were observed after lunch breaks (day and weekend shifts) and with the arrival of a rested resident during overlapping on-call shifts.[67]		Yes	
12 – Infection prevention related			**2**	**0**	**1**
Hand hygiene compliance	Time of day	The hours from 12:00–14:00 had a notably lower compliance than other periods of the day.[16]		Unclear	
	Time into shift	Hand hygiene compliance rates dropped by from the beginning to the end of a typical 12-hr work shift.[18]		Yes	
	Hours off since previous shift	Longer breaks between work shifts increased subsequent compliance rates.[18]		Yes	

1 Addis (2020), 2 Agarwal (2022), 3 Allan (2019), 4 Almadi (2015), 5 Anastasian (2014), 6 Backmann (2021b), 7 Backmann (2021a), 8 Beg (2021), 9 Bernstein (2022), 10 Bersani (2020), 11 Bonafide (2017), 12 Canfield (2020), 13 Castrén (2015), 14 Chan (2009), 15 Changolkar (2020), 16 Cheng (2011), 17 Christensen (1977), 18 Dai (2015), 19 Delaney (2005), 20 Dexheimer (2017), 21 Dominello (2015), 22 Edwards (2003), 23 Ergün-Sahin (2022), 24 Evans (1984), 25 Faulk (2013), 26 Freedman (2011), 27 Ginestra (2023), 28 Hare (2021), 29 Harewood (2009), 30 Hsiang (2019), 31 Hueston (1996), 32 Hughes (2020), 33 Hunt (2021), 34 Kaneshiro (2010), 35 Keswani (2016), 36 Kim (2018), 37 Kolla (2023), 38 Kushnir (2020), 39 Lee (2017), 40 Leffler (2012), 41 Linder (2014), 42 Long (2011), 43 Lu (2023), 44 Lurix (2012), 45 Marcondes (2018), 46 McClelland (2013), 47 Mitchell (2006), 48 Munson (2011), 49 Neprash (2019), 50 Oakes (2021), 51 Paeck (2013), 52 Persson (2019), 53 Philpot (2018), 54 Pignatiello (2022), 55 Pu (2019), 56 Rawshani (2017), 57 Saposnik (2017), 58 Soukup (2019a), 59 Soukup (2020a), 60 Soukup (2020b), 61 Sapci (2022), 62 Stecker (2015), 63 Taylor-Phillips (2015), 64 Taylor-Phillips (2016), 65 Thurtle (2014), 66 Trinh (2021), 67 Vosshenrich (2021), 68 Wright (2006), 69 Xu (2018), 70 Zhang (2023a), 71 Zhang (2023b), 72 Zheng (2020).

The probability of any diagnostic test order (type of test not specified further) did not change significantly with every additional patient that was examined and neither did the probability of ordering a lab test (type of test not specified further) (Ergün-Şahin et al., 2022). Ordering rates of CT chest scans did not change significantly over the shift (Zheng et al., 2020).

DC2: Evaluating Test Results. The decision fatigue hypothesis was examined in 29 studies (Almadi et al., 2015; Backmann et al., 2021b, 2021a; Beg et al., 2021; Bernstein et al., 2022; Bonafide et al., 2017; Chan et al., 2009; Christensen et al., 1977; Freedman et al., 2011; Harewood et al., 2009; Kaneshiro et al., 2010; Keswani et al., 2016; A. Lee et al., 2017; Leffler et al., 2012; Long et al., 2011; Lu et al., 2023; Lurix et al., 2012; Marcondes et al., 2018; Mitchell et al., 2006; Munson et al., 2011; Paeck et al., 2013; Pu et al., 2019; Sapci et al., 2022; Taylor-Phillips et al., 2015, 2016; Thurtle et al., 2014; Xu et al., 2018; Zhang, Chen, Cao, et al., 2023; Zhang, Chen, Wang, et al., 2023b) including 51 formal tests of decision fatigue. Twenty-one analyses supported the hypothesis (Almadi et al., 2015; Backmann et al., 2021b; Beg et al., 2021; Bernstein et al., 2022; Bonafide et al., 2017; Chan et al., 2009; Kaneshiro et al., 2010; Leffler et al., 2012; Long et al., 2011; Lu et al., 2023; Marcondes et al., 2018; Paeck et al., 2013; Pu et al., 2019; Taylor-Phillips et al., 2016; Xu et al., 2018; Zhang, Chen, Cao, et al., 2023; Zhang, Chen, Wang, et al., 2023b), 11 remained inconclusive (Backmann et al., 2021b, 2021a; Christensen et al., 1977; Leffler et al., 2012; Long et al., 2011; Lu et al., 2023; Munson et al., 2011; Taylor-Phillips et al., 2015; Xu et al., 2018), and 19 did not support the hypothesis (Backmann et al., 2021b; Beg et al., 2021; Bernstein et al., 2022; Freedman et al., 2011; Harewood et al., 2009; Keswani et al., 2016; A. Lee et al., 2017; Leffler et al., 2012; Lurix et al., 2012; Mitchell et al., 2006; Paeck et al., 2013; Sapci et al., 2022; Taylor-Phillips et al., 2015, 2016; Thurtle et al., 2014).

The studies that supported the hypothesis showed that the adenoma detection rate via endoscopy significantly decreased with increasing time into the shift (Almadi et al., 2015; Paeck et al., 2013), increasing numbers of decisions made previously (Almadi et al., 2015; Leffler et al., 2012), increasing time of day (Kaneshiro et al., 2010; Lu et al., 2023; Pu et al., 2019; Xu et al., 2018), and increasing order number of decision (Marcondes et al., 2018). Polyp detection rate via endoscopy also significantly decreased with increasing time of day (Chan et al., 2009; Kaneshiro et al., 2010; Zhang, Chen, Cao, et al., 2023), and increasing time of day regardless of shift timing or duration (Long et al., 2011). Furthermore, lesion detection via endoscopy significantly decreased as the day progressed and so did the endoscopy biopsy rate (Zhang, Chen, Wang, et al., 2023b). Endoscopy reading time decreased with increasing number of decisions made previously (Lu et al., 2023). For mammography reading decisions, the predicted sensitivity (true positive rate) gradually decreased throughout the day (Backmann et al., 2021b) and the odds of false-positive findings and patients being recalled increased (Bernstein et al., 2022). Another study found that the rate of patient recall decreased with increasing order number of decision (Taylor-Phillips et al., 2016). With each successive hour that passed in a nurse's shift, response time to physiologic monitor alarms was significantly slower (Bonafide et al., 2017).

Results were inconclusive for the following cases: Xu et al. found that the adenoma detection rate via colonoscopy increased over time, peaking between 09:00–10:00 and again after the lunch break (Xu et al., 2018). Conversely, progression of time was significantly associated with reduced odds of adenoma detection, however, reductions with regards to shift timing or duration did not reach statistical significance (Long et al., 2011). For polyp detection via colonoscopy, Munson et al. showed that the polyp detection rate did not decline steadily throughout the day, it declined during the afternoon shift (13:30 - 16:30) but remained stable otherwise (Munson et al., 2011) and similarly Lu at al. found no clear trend in the way the polyp detection rate changed over time (Lu et al., 2023). Looking at the number of decisions made previously, authors found no significant variation in polyp detection rate, however in post hoc analysis, polyp detection rates were significantly lower after the fifth case of the day for endoscopists with low volumes of cases and after the tenth case of the day for endoscopists with high volumes of cases (Leffler et al., 2012). For mammography reading decisions, true and false positives were negatively associated with image position within a batch (Backmann et al., 2021a), and rates of true and false negatives were positively associated with image position within a batch (the second was not statistically significant (Backmann et al., 2021a)). For time of day, the percentage of positive readings was highest before lunch and in the early afternoon (Backmann et al., 2021b). Finally, there was a sensitivity (true positive rate) decrease and specificity (true negative rate) increase over the course of reading 100 chest x-rays (Taylor-Phillips et al., 2015), 60 bone fracture x-rays (Taylor-Phillips et al., 2015), and 100 chest CT scans (Taylor-Phillips et al., 2015). Results were also inconclusive for radiograph reading decisions, more specifically the detection of pulmonary nodules, where for some residents, detection performance improved in the fatigued state, while for others, performance declined (Christensen et al., 1977).

Lee et al. and Leffler et al. found no significant inverse relationship between adenoma detection rates via colonoscopy and the number of decisions made previously (A. Lee et al., 2017; Leffler et al., 2012). This was also the case for adenoma detection rate and time of day (Keswani et al., 2016; Leffler et al., 2012; Lurix et al., 2012; Paeck et al., 2013; Sapci et al., 2022), and no decrease in adenoma detection rate was noted by Lurix et al. with an increasing queue position (Lurix et al., 2012). Similarly, the polyp/lesion detection rate was not found to be significantly related to time of day (Freedman et al., 2011; Leffler et al., 2012; Paeck et al., 2013), or the order number of decision (Beg et al., 2021; Harewood et al., 2009; Thurtle et al., 2014). Backmann et al. found a gradually increasing predicted specificity (true negative rate) throughout the day (Backmann et al., 2021b), and that breast cancer detection rate for individual readers did not change with the order number of decisions, as represented by near identical odds of detecting cancer between the first and 40th cases (Taylor-Phillips et al., 2015, 2016). Furthermore, Bernstein et al. found that the odds of a true positive finding when reading mammography results increased with every additional hour of reading time (Bernstein et al., 2022). For cervical cytology test interpretation decisions, there were no statistically significant changes in false negative rates throughout the day (Mitchell et al., 2006).

DC3 – Defining a Problem. The decision fatigue hypothesis was examined in seven studies (Edwards et al., 2003; Kushnir et al., 2020; Soukup, Gandamihardja, et al., 2019; Soukup, Lamb, Morbi, et al., 2020; Soukup, Lamb, Shah, et al., 2020b; Stecker et al., 2015; Trinh et al., 2021) including 20 cases of formal tests of decision fatigue. Nine cases supported the hypothesis (Soukup, Gandamihardja, et al., 2019; Soukup, Lamb, Morbi, et al., 2020; Soukup, Lamb, Shah, et al., 2020b; Stecker et al., 2015; Trinh et al., 2021), two remained inconclusive (Edwards et al., 2003; Kushnir et al., 2020), and nine did not support the hypothesis (Kushnir et al., 2020; Soukup, Lamb, Morbi, et al., 2020; Soukup, Lamb, Shah, et al., 2020b; Stecker et al., 2015).

Multidisciplinary oncology team decision-making for cancer patients significantly declined in quality for certain indicators with the order number of decisions. This was in terms of the quality of the presented patient information (Soukup, Gandamihardja, et al., 2019; Soukup, Lamb, Morbi, et al., 2020), and the quality of disciplinary contribution (Soukup, Gandamihardja, et al., 2019; Soukup, Lamb, Morbi, et al., 2020). There was also a negative correlation with asking questions and positive reactions (Soukup, Lamb, Morbi, et al., 2020), showing solidarity, and agreeing (Soukup, Lamb, Shah, et al., 2020b), and asking for orientation and opinion (Soukup, Lamb, Shah, et al., 2020b). The serial position of the treatment decision was positively correlated with disagreeing (Soukup, Lamb, Shah, et al., 2020b). Outwith the multidisciplinary cancer team domain, Trinh et al. found that some physicians assessed a significantly different number of diagnoses in the first vs last hour of their shift (Trinh et al., 2021).

Results were inconclusive for temperature recording decisions (Edwards et al., 2003). Temperature recording demonstrated a strong temporal pattern with temperatures recorded more frequently between 19:00 and 05:00 and less frequently between 7:00 and 17:00 (Edwards et al., 2003). Decisions linked to the use of the Finnegan neonatal scoring tool showed that the mean high NAS scores decreased as the time of day increased, and daily high NAS scores tended to occur mostly between 12:00 and 15:00, which is in the middle of the observer's shift (Kushnir et al., 2020). Neither of these time trends clearly supports or opposes the decision fatigue hypothesis.

The following results showed no support for the decision fatigue hypothesis: there was no relationship between the scoring observer's shift hour and daily high NAS scores (Kushnir et al., 2020), and when comparing the calculated NAS score using only objective data points to the total score, no significant difference occurred throughout the day (Kushnir et al., 2020). Multidisciplinary team decision-making for cancer patients showed no change in quality for certain indicators with the order number of decisions in terms of giving answers and negative reactions (Soukup, Lamb, Morbi, et al., 2020) and showing antagonism (Soukup, Lamb, Shah, et al., 2020b) and giving orientation and suggestions (Soukup, Lamb, Shah, et al., 2020b). Finally, there was no effect of the time of day on the rate at which the diagnosis of stroke was made (Stecker et al., 2015). Further quality indicators of multidisciplinary team decision-making for cancer patients showed correlations with the order number of decisions that were in opposing directions to the decision fatigue hypothesis: showing tension was negatively correlated (Soukup, Lamb, Morbi, et al., 2020), and tension release (Soukup, Lamb, Morbi, et al., 2020) and giving opinions (Soukup, Lamb, Morbi, et al., 2020) were positively correlated with the order number of decisions.

DC4 – Drug related. The decision fatigue hypothesis was examined in 12 studies (Agarwal et al., 2022; Dexheimer et al., 2017; Edwards et al., 2003; Hare et al., 2021; Hughes et al., 2020; Linder et al., 2014; Neprash & Barnett, 2019; Oakes et al., 2021; Philpot et al., 2018; Saposnik et al., 2017; Stecker et al., 2015) including 15 cases of formal tests of decision fatigue. Seven cases supported the hypothesis (Hare et al., 2021; Hughes et al., 2020; Linder et al., 2014; Neprash & Barnett, 2019; Oakes et al., 2021; Philpot et al., 2018; Saposnik et al., 2017), five remained inconclusive (Agarwal et al., 2022; Dexheimer et al., 2017; Edwards et al., 2003; Ginestra et al., 2023; Neprash & Barnett, 2019), and three did not support the hypothesis (Neprash & Barnett, 2019; Stecker et al., 2015).

The likelihood of statin prescribing significantly decreased (Hare et al., 2021; Oakes et al., 2021) and the likelihood of antibiotic prescribing (Linder et al., 2014) and opioid prescribing (Hughes et al., 2020; Neprash & Barnett, 2019; Philpot et al., 2018) significantly increased throughout the day. Therapeutic decisions for hypothetical MS patients saw a significant increase in the prevalence of therapeutic inertia in the second set of case scenarios compared to an initial set (Saposnik et al., 2017).

Results were inconclusive for decisions to respond to medication order alerts (Dexheimer et al., 2017), and antipyretic administration (Edwards et al., 2003). For both, significant changes across the day were observed, however, the nature of those changes was not clearly (contra-)indicative of the decision fatigue hypothesis. Neprash et al. found that initial prescriptions for statins did not increase as a function of appointment order (Neprash & Barnett, 2019), but they did not report whether statin prescribing significantly decreased, which would support the hypothesis. Agarwahl et al. found that among visits for which opioids were sometimes indicated, prescribing increased per hour on shift, however among visits for which opioids were usually not indicated, opioid prescriptions fell per hour on shift (Agarwal et al., 2022).

There was no association between appointment timing and NSAID prescribing (Neprash & Barnett, 2019), initial orders for antihypertensives did not increase as a function of appointment order (Neprash & Barnett, 2019), and there was no effect of the time of day on the rate at which tPA (tissue plasminogen activator) was given (Stecker et al., 2015).

DC5 – Therapeutic procedure related. The decision fatigue hypothesis was examined in 12 studies (Addis et al., 2020; Anastasian et al., 2014; Changolkar et al., 2020; Delaney & Fogg, 2005; Dominello et al., 2015; Evans et al., 1984; Harewood et al., 2009; Hueston et al., 1996; Kim et al., 2018; Persson et al., 2019; Stecker et al., 2015; Wright et al., 2006) including 16 cases of formal tests of decision fatigue. Six cases supported the hypothesis (Addis et al., 2020; Anastasian et al., 2014; Changolkar et al., 2020; Kim et al., 2018; Persson et al., 2019; Wright et al., 2006), six remained inconclusive (Addis et al., 2020; Delaney & Fogg, 2005; Evans et al., 1984; Harewood et al., 2009; Hueston et al., 1996; Stecker et al., 2015), and four did not support the hypothesis (Addis et al., 2020; Dominello et al., 2015).

In line with the decision fatigue hypothesis, the probability of giving an intraoperative red blood cell transfusion significantly increased with later case start times (Addis et al., 2020), case end time significantly correlated with delayed extubation in multilevel spine surgery (Anastasian et al., 2014), and appointment time of day was associated with a significant decrease in influenza vaccination rates (Changolkar et al., 2020; Kim et al., 2018). Patients who met a surgeon toward the end of their shift were significantly less likely to be scheduled for an operation compared with those who were seen first (Persson et al., 2019) and the predicted probability for decisions that resulted in anaesthetic adverse events increased from a low of 1.0% at 09:00 to a high of 4.2% at 16:00 (Wright et al., 2006).

Results were inconclusive for decisions about giving any blood product (red blood cells, plasma, platelets, cryoprecipitate) transfusion during cardiac surgery (Addis et al., 2020) as there were significant changes with later case start times, however, the direction of those changes was not in clear support of the decision fatigue hypothesis. Decisions about delivery via caesarean section (vs vaginal delivery) changed with time of day, but again those changes were not clearly (contra-)indicative of the decision fatigue hypothesis (Evans et al., 1984; Hueston et al., 1996). Cecal intubation, which is the successful completion of colonoscopy, saw a decline with successive procedures; however, this was only the case for morning procedures and not for afternoon procedures (Harewood et al., 2009). Finally, there was evidence that the frequency of stroke alerts varied throughout the day; however, the direction of those changes was again not in clear support of the hypothesis (Stecker et al., 2015).

There was no relationship between the time of day and the probability of interoperative transfusion of plasma (Addis et al., 2020), cryoprecipitate (Addis et al., 2020), or platelets (Addis et al., 2020). For radiation treatment delivery errors, there was also no significant difference in error rate based on the time of day (Dominello et al., 2015).

DC7 – Contact related. The decision fatigue hypothesis was examined in six studies (Allan et al., 2019; Castrén et al., 2015; Faulk et al., 2013; Neprash & Barnett, 2019; Rawshani et al., 2017; Zheng et al., 2020) including nine cases of formal tests of decision fatigue. Three cases supported the hypothesis (Allan et al., 2019; Faulk et al., 2013), two remained inconclusive (Castrén et al., 2015; Rawshani et al., 2017), and four did not support the hypothesis (Allan et al., 2019; Neprash & Barnett, 2019; Zheng et al., 2020).

Nurses' triaging decisions via a medical telephone helpline were significantly more likely to be conservative (referral within 12-hr vs outside 12-hr window), for every additional call taken since nurses' last break (Allan et al., 2019), and for every hour that passed since the last break (Allan et al., 2019). Patients admitted to an inpatient rehabilitation facility later in the day were significantly more likely to be returned to the acute care hospital than patients admitted earlier in the day (Faulk et al., 2013).

Results were inconclusive for pre-hospital assessment decisions: the frequency of patients with non-specific complaints peaked between approximately 10:00 and 14:00, while the frequency of patients with specific conditions had its peak at approximately 09:00 (Castrén et al., 2015) a trend not clearly (contra-)indicative of the decision fatigue hypothesis. For assigning case priority, the probability of not giving priority 1 when indicated increased during the morning and reached a maximum at 12:59, however, the risk declined thereafter (Rawshani et al., 2017), which is not in clear support of the hypothesis.

In comparison to the effects found in regards to time passed since the last rest break, there was no effect of the number of decisions (Allan et al., 2019), or time elapsed (Allan et al., 2019) since the start of the shift on triage decisions. The ordering of referral to physical therapy did not significantly change as the day progressed (Neprash & Barnett, 2019). Finally, disposition decisions (whether a patient was discharged or admitted) did not change significantly with time into the shift (Zheng et al., 2020).

DC9 – Treatment goal. The decision fatigue hypothesis was examined in one study (Kolla et al., 2023) which included one formal test of decision fatigue that supported the decision fatigue hypothesis.

Oncology clinicians' likelihood of having and documenting serious illness conversations decreased as their clinic sessions progressed (Kolla et al., 2023).

DC11 – Documentation related. The decision fatigue hypothesis was examined in two studies (Canfield et al., 2020; Vosshenrich et al., 2021) including three cases of formal tests of decision fatigue. One case supported the hypothesis (Vosshenrich et al., 2021), one remained inconclusive (Canfield et al., 2020), and one did not support the hypothesis (Canfield et al., 2020).

The accuracy of written radiology reports decreased significantly with increasing resident work hours (Vosshenrich et al., 2021).

Results were inconclusive for patient order entry decisions into electronic health records; time of day appeared to be a factor, with wrong patient order entry rates being slightly lower during the night than the day, however, this trend is not clearly (contra-)indicative of the decision fatigue hypothesis (Canfield et al., 2020).

There was no clear pattern of order retraction of patient orders from electronic health records by hour (Canfield et al., 2020).

DC12 – Infection prevention related. The decision fatigue hypothesis was examined in two studies (Cheng et al., 2011; Dai et al., 2015), including three cases of formal tests of decision fatigue. Two cases supported the hypothesis (Dai et al., 2015) and one remained inconclusive (Cheng et al., 2011).

Hand hygiene compliance rates reduced significantly from the beginning to the end of a typical 12-hr work shift (Dai et al., 2015) and longer breaks between work shifts significantly increased subsequent compliance rates (Dai et al., 2015).

Cheng et al. found that the hours from 12:00–14:00 had a notably lower hand hygiene compliance than other periods of the day, which is not in clear support of the decision fatigue hypothesis (Cheng et al., 2011).

Qualitative evidence, evidence from expert opinions, editorials, book chapters, and literature reviews

In the qualitative study, decision fatigue was not addressed as the main study question but emerged as a theme in the context of the perceived benefits and drawbacks of 24-hour on-call shifts (Cygler et al., 2021). Individual interviews and focus groups were conducted with 17 internal medicine residents, to understand what residents perceived as the benefits and detriments of 24-hour, in-house (at the hospital) call. Seven themes were generated, one of them being multidimensional fatigue related to call duration, including decision fatigue, emotional fragility and lability, and loss of empathy. Residents acknowledged that they were affected by decision fatigue. They described 'perceived cognitive dysfunction, both toward the end of on-call periods and on the following day, which manifested as slowed processing time, deficits in working memory, and overall impaired executive functioning due to the heavy cognitive load. Residents attributed this cognitive load to both sleep deprivation and the consecutive number of hours spent working' (Cygler et al., 2021).

The first opinion piece focused on decision fatigue in primary care clinicians (Moorhouse, 2020). The author was a trainee in primary care and described high patient load, increasingly complex patients, and the lack of dedicated breaks as the main drivers of decision fatigue in this setting. Moorhouse (2020) suggested this not only leads to poor clinical decision-making, and decision avoidance but can also act as a precursor to burnout.

The second opinion piece listed decision fatigue as one of the main explanations for why quality of care has been reported to decrease over the course of the day (Oakes & Patel, 2021). The authors stated that a combination of shared cognitive biases, hurried visits, and mounting decision fatigue were the reasons why the probability of making high-value care decisions decreased later in the day, and the probability of making low-value care decisions increased. Rather than referring to their own experiences, the authors pulled together examples from the literature to discuss the phenomenon.

In the third opinion piece, Schweitzer et al. mainly focused on decision fatigue and limited willpower in the broad healthcare context (Schweitzer et al., 2023). The authors introduced a conceptual framework to visualise the theoretical connection between internal resources, external resources, willpower, self-control, and decision fatigue in clinicians. Three case studies about physicians, nurses, and physiotherapists in a hospital setting were described to illustrate that depleted will-power can compromise healthcare delivery. Research areas for addressing decision fatigue in the healthcare context were proposed.

The first editorial focused on decision fatigue in the emergency department (Dubash et al., 2020). The authors were two Emergency Medicine Advanced Trainees and one Emergency Physician, who drew upon their expertise and experiences when describing factors in their work setting that contribute to decision fatigue and exemplify effects of decision fatigue. They listed large decision volume, multi-tasking, interruptions, and the unscheduled nature of work alongside a list of individual and departmental factors as contributors to decision fatigue. They further suggested that decision fatigue can manifest in different ways and may not always be evident to the individual. They suggested that manifestations include basic cognitive errors such as 'confusing patients with similar presentations, decision procrastination, making arbitrary choices or ones that are the 'easy option' and without consideration of long-term consequences. At the extreme end, one may demonstrate impulsive and unsafe behaviour 'resembling someone who's been drinking'' (Dubash et al., 2020).

The second editorial focused on decision fatigue in nurses, specifically in the context of the COVID-19 pandemic (Hatami et al., 2022). The authors discussed how the pandemic put additional strains on nurses, both in terms of physical and mental workload. The article lacks a clear aim and makes generic statements relating to decision fatigue in a pandemic situation.

The third editorial focused on decision fatigue in the emergency department in hospitals (Nasa & Majeed, 2023). Associated factors such as working hours, workload, work schedule, lack of resources, multitasking, and burnout were discussed. Effects of decision fatigue on patient care but also on personal life were listed. Mitigating strategies were discussed, including early recognition of decision fatigue through validated screening tools, behavioural education via simulation-based training to simplify decision-making, and developing a priority system to defer less critical decisions.

The narrative review described a selection of cognitive and system-based sources of detection and interpretation errors in diagnostic radiology and lists 'Decision (Mental) Fatigue' as one of them (C. S. Lee et al., 2013). The authors suggested that radiology residents who provide preliminary interpretations independently during off hours are especially vulnerable to decision fatigue, and those working prolonged shifts, off hours and with high-volume or high-complexity tasks are at the greatest risk.

The book chapter specifically focused on decision fatigue in clinicians in the hospital emergency department (X. Q. Lee, 2021). Four case studies were presented that describe situations where clinicians become decision fatigued. The effects of this decision fatigue were described as poor communication patterns, choosing the easy option, avoiding difficult conversations, oversimplification, anchoring to what seemed the obvious diagnosis, and ordering unnecessary diagnostic imaging. A list of potential solutions was provided, including managing contributing factors like hunger and tiredness, performing regular self-checks, taking breaks, prioritising major decisions early, sharing the cognitive load by delegating, avoiding procrastination, using algorithms and checklists, and implementing organisational interventions like limiting shift durations and improving rostering practices.

The literature review focused on understanding decision-making in multidisciplinary oncology team meetings. It did not use the term 'decision fatigue', however, the review's focus was on how time-on-task affects decision-making, which matches our inclusion criteria (Soukup, Lamb, et al., 2019b). The authors identified seven cognitive–behavioural pitfalls resulting from time-on-task, which they suggested have potential quality and safety implications for multi-disciplinary team meetings. They listed reduced rational thinking, impulsive and riskier decisions, reduced attention, reduced ability to monitor and detect errors, reduced task persistence, reduced trust and morality and status quo and effort avoidance as the cognitive–behavioural pitfalls associated with prolonged periods of cognitive activity (Soukup, Lamb, et al., 2019b).

Theoretical conceptualisation of decision fatigue

Among the 82 articles reviewed, decision fatigue was explicitly named as the primary research focus in 20 articles (Agarwal et al., 2022; Allan et al., 2019; Dubash et al., 2020; Hatami et al., 2022; Hsiang et al., 2019; Hughes et al., 2020; Hunt et al., 2021; Kim et al., 2018; Kolla et al., 2023; X. Q. Lee, 2021; Linder et al., 2014; Moorhouse, 2020; Nasa & Majeed, 2023; Oakes & Patel, 2021; Persson et al., 2019; Pignatiello et al., 2022; Schweitzer et al., 2023; Soukup, Gandamihardja, et al., 2019; Trinh et al., 2021; Zheng et al., 2020). In those articles, decision fatigue was named in the introduction and further elaboration on decision fatigue was provided in the discussion section. Six articles introduced the decision fatigue concept in their discussion sections (Addis et al., 2020; Ginestra et al., 2023; Hare et al., 2021; Oakes et al., 2021; Philpot et al., 2018; Saposnik et al., 2017), as a potential explanation for their findings but did not place a specific focus on the concept throughout the study. One of these studies also mentioned decision fatigue in their methods section (Saposnik et al., 2017). Two articles, comprising one review and one qualitative study, referenced decision fatigue within their results sections (Cygler et al., 2021; C. S. Lee et al., 2013), while a singular article made a solitary mention of decision fatigue in its appendix (Soukup, Lamb, et al., 2019b). The term decision fatigue was not explicitly mentioned in 53 studies (Almadi et al., 2015; Anastasian et al., 2014; Backmann et al., 2021b, 2021a; Beg et al., 2021; Bernstein et al., 2022; Bersani et al., 2020; Bonafide et al., 2017; Canfield et al., 2020; Castrén et al., 2015; Chan et al., 2009; Changolkar et al., 2020; Cheng et al., 2011; Christensen et al., 1977; Dai et al., 2015; Delaney & Fogg, 2005; Dexheimer et al., 2017; Dominello et al., 2015; Edwards et al., 2003; Ergün-Şahin et al., 2022; Evans et al., 1984; Faulk et al., 2013; Freedman et al., 2011; Harewood et al., 2009; Hueston et al., 1996; Kaneshiro et al., 2010; Keswani et al., 2016; Kushnir et al., 2020; A. Lee et al., 2017; Leffler et al., 2012; Long et al., 2011; Lu et al., 2023; Lurix et al., 2012; Marcondes et al., 2018; McClelland et al., 2013; Mitchell et al., 2006; Munson et al., 2011; Neprash & Barnett, 2019; Paeck et al., 2013; Pu et al., 2019; Rawshani et al., 2017; Sapci et al., 2022; Soukup, Gandamihardja, et al., 2019; Soukup, Lamb, Morbi, et al., 2020; Soukup, Lamb, Shah, et al., 2020b; Stecker et al., 2015; Taylor-Phillips et al., 2015, 2016; Thurtle et al., 2014; Vosshenrich et al., 2021; Wright et al., 2006; Xu et al., 2018; Zhang, Chen, Cao, et al., 2023; Zhang, Chen, Wang, et al., 2023b). Among the articles where the term was not explicitly mentioned, various alternate terms such as 'operator fatigue' (Almadi et al., 2015; Harewood et al., 2009; Long et al., 2011), 'provider fatigue' (Anastasian et al., 2014), 'colonoscopist fatigue' (Freedman et al., 2011), 'endoscopist fatigue' (A. Lee et al., 2017; Lurix et al., 2012; Munson et al., 2011; Thurtle et al., 2014; Xu et al., 2018), 'physician fatigue' (Paeck et al., 2013), 'fatigue' (Backmann et al., 2021a; Beg et al., 2021; Bernstein et al., 2022; Christensen et al., 1977; Faulk et al., 2013; Hueston et al., 1996; Kaneshiro et al., 2010; Keswani et al., 2016; Kushnir et al., 2020; Lu et al., 2023; Marcondes et al., 2018; McClelland et al., 2013; Pu et al., 2019; Taylor-Phillips et al., 2015, 2016; Vosshenrich et al., 2021), 'circadian rhythms/variations/lows' (Backmann et al., 2021b; Rawshani et al., 2017; Wright et al., 2006), and 'alert fatigue' (Dexheimer et al., 2017) were employed to describe similar concepts.

Definitions of decision fatigue

A total of 27 records provided definitions of the decision fatigue phenomenon (Addis et al., 2020; Agarwal et al., 2022; Allan et al., 2019; Dubash et al., 2020; Ginestra et al., 2023; Hare et al., 2021; Hatami et al., 2022; Hsiang et al., 2019; Hughes et al., 2020; Hunt et al., 2021; Kim et al., 2018; Kolla et al., 2023; X. Q. Lee, 2021; Linder et al., 2014; Moorhouse, 2020; Nasa & Majeed, 2023; Oakes et al., 2021; Oakes & Patel, 2021; Persson et al., 2019; Pignatiello et al., 2022; Saposnik et al., 2017; Schweitzer et al., 2023; Soukup, Gandamihardja, et al., 2019; Soukup, Lamb, et al., 2019b; Trinh et al., 2021; Zheng et al., 2020). Two articles that used the term decision fatigue did not provide a definition (Cygler et al., 2021; Philpot et al., 2018). Nine of the definitions described decision fatigue in association with the erosion/loss/depletion of self-control due to repeated decision-making instances (Addis et al., 2020; Agarwal et al., 2022; Hsiang et al., 2019; Kim et al., 2018; Linder et al., 2014; Oakes et al., 2021; Oakes & Patel, 2021; Schweitzer et al., 2023; Trinh et al., 2021). Others referred to a depletion of mental resources (Allan et al., 2019; Hatami et al., 2022; Hughes et al., 2020), cognitive resources (Hunt et al., 2021; Nasa & Majeed, 2023), or psychological resources (Dubash et al., 2020). One definition pointed towards the decline of motivation to exert mental effort over time (Allan et al., 2019). Many of the definitions included potential outcomes of decision fatigue such as a shift toward easier, safer, or more gratifying decisions and actions (Allan et al., 2019), favouring the status quo or the cognitively 'easier' choice (Hughes et al., 2020), a decrement in the ability to make appropriate/optimal decisions (Ginestra et al., 2023; Soukup, Lamb, et al., 2019b), the inability to continue making difficult decisions (Kolla et al., 2023), the deterioration in the quality of those decisions (Dubash et al., 2020; Hatami et al., 2022; X. Q. Lee, 2021), the erosion/depletion of (active) initiative (Hare et al., 2021; Hsiang et al., 2019; Kim et al., 2018; Oakes et al., 2021), or the reduced ability to make decisions and regulate behaviour (Trinh et al., 2021). Others remained vague and noted simply that decision fatigue negatively affects subsequent decisions, leading to performance decrements over time (Soukup, Gandamihardja, et al., 2019), or influences subsequent decisions (Zheng et al., 2020) without defining the nature of influence.

Theoretical frameworks used in explaining decision fatigue

Out of the 82 included records, seven cited one or more theoretical frameworks as an explanation of the decision fatigue phenomenon (Allan et al., 2019; Hatami et al., 2022; Persson et al., 2019; Pignatiello et al., 2022; Schweitzer et al., 2023; Soukup, Lamb, et al., 2019b; Trinh et al., 2021).

All seven (Allan et al., 2019; Hatami et al., 2022; Persson et al., 2019; Pignatiello et al., 2022; Schweitzer et al., 2023; Soukup, Lamb, et al., 2019b; Trinh et al., 2021) mentioned the **Strength/ Resource Model of Self-Control** (Baumeister, 2003; Baumeister et al., 2007), a theoretical framework which proposes that mental/cognitive resources are limited/finite and become depleted when self-control is exercised. This theoretical framework implies that continuous decision-making requires self-control, therefore self-control becomes depleted during episodes of continuous decision-making and when the limited resources have been depleted there is a decline in decision quality or a shift towards easier decisions. Sixteen articles alluded to the strength/resource model of self-control without naming it directly by referring to decision fatigue as the 'depletion of self-control' (also referred to as 'ego depletion') (Addis et al., 2020; Agarwal et al., 2022; Dubash et al., 2020; Ginestra et al., 2023; Hare et al., 2021; Hsiang et al., 2019; Hughes et al., 2020; Hunt et al., 2021; Kim et al., 2018; X. Q. Lee, 2021; Linder et al., 2014; Moorhouse, 2020; Nasa & Majeed, 2023; Oakes et al., 2021; Oakes & Patel, 2021; Soukup, Gandamihardja, et al., 2019).

Three of the studies (Allan et al., 2019; Soukup, Lamb, et al., 2019b; Trinh et al., 2021) also mentioned the **Process Model of Self-Control** (Inzlicht et al., 2014; Inzlicht & Schmeichel, 2012), a theoretical framework which highlights shifting attention and motivation as contributors to regulatory failures. Rather than suggesting that self-control is a finite resource that can be depleted, this model suggests that shifts in motivation lead to the switching of task priorities.

One study (Soukup, Lamb, et al., 2019b) mentioned the **Opportunity cost model of subjective effort and task performance** (Kurzban et al., 2013), a framework which suggests that the aversive experience of mental effort and subsequent performance deterioration arises from limited computational resources associated with executive function. The limited resources lead to a trade-off between current task demands and alternative cognitive uses, ultimately influencing effort allocation and performance outcomes.

Articles that did not explicitly mention decision fatigue (Backmann et al., 2021a; McClelland et al., 2013; Taylor-Phillips et al., 2015, 2016) link the phenomenon to theories relating to vigilance decrement (Mackworth, 1948; See et al., 1995), and the controlled attention model (Kane & Engle, 2002).

Methodological quality of included studies

The quality ratings are summarised in Table 3.

Table 3. Methodological quality ratings of included studies.

Quantitative & Qualitative Studies (using the Mixed Methods Appraisal Tool (MMAT))					
Study (Author & Year)	**Methodological quality criteria** (dependent on Study Category)				
Qualitative Studies					
Methodological quality criteria for qualitative studies	Is the qualitative approach appropriate to answer the research question?	Are the qualitative data collection methods adequate to address the research question?	Are the findings adequately derived from the data?	Is the interpretation of results sufficiently substantiated by data?	Is there coherence between qualitative data sources, collection, analysis and interpretation?
Cygler 2021	Yes	Yes	Yes	Yes	Yes
Randomized Controlled Trials					
Methodological quality criteria for quantitative randomised controlled trials	Is randomisation appropriately performed?	Are the groups comparable at baseline?	Are there complete outcome data?	Are outcome assessors blinded to the intervention provided?	Did the participants adhere to the assigned intervention?
Bersani 2020	Can't tell	Can't tell	Yes	No	Yes
Saposnik 2017	Yes	Yes	Yes	Yes	Can't tell
Taylor-Phillips 2016	Yes	Yes	Yes	Yes	Yes
Non-Randomised Studies					
Methodological quality criteria for quantitative non-randomised studies	Are the participants representative of the target population?	Are measurements appropriate regarding both the outcome and intervention (or exposure)?	Are there complete outcome data?	Are the confounders accounted for in the design and analysis?	During the study period, is the intervention administered (or exposure occurred) as intended?
Addis 2020	Can't tell	Can't tell	Yes	Yes	Yes
Agarwal 2022	Can't tell	Yes	Yes	Yes	Yes
Allan 2019	Yes	Yes	Yes	No	Yes
Almadi 2015	Yes	Yes	Yes	Yes	Yes
Anastasian 2014	Can't tell	No	Yes	Yes	Yes
Backmann 2021a	Yes	Yes	Yes	Yes	Yes
Backmann 2021b	Yes	Yes	Yes	Yes	Yes
Beg 2021	No	Yes	Yes	No	Yes
Bernstein 2022	Yes	Yes	Yes	Yes	Yes
Bonafide 2017	Can't tell	Yes	No	Yes	Yes
Canfield 2020	Can't tell	No	Can't tell	Yes	Yes
Castrén 2015	Can't tell	Yes	Yes	Yes	Yes
Chan 2009	Can't tell	Yes	Yes	Yes	Yes
Changolkar 2020	Can't tell	Yes	Yes	Yes	Yes
Cheng 2011	Can't tell	Yes	No	No	Yes
Christensen 1977	No	Yes	Yes	No	Yes
Dai 2015	Can't tell	Yes	Yes	Yes	Yes
Delaney 2005	Can't tell	Can't tell	Yes	No	Yes
Dexheimer 2017	Yes	Yes	Yes	Can't tell	Yes
Dominello 2015	Yes	Yes	Yes	No	Yes
Edwards 2003	Can't tell	Yes	Can't tell	No	Yes
Ergün-Sahin 2022	Can't tell	Yes	Can't tell	Yes	Yes

Evans 1984	Can't tell	Yes	No	No	Yes
Faulk 2013	Can't tell	Yes	Yes	Yes	Yes
Freedman 2011	Can't tell	Yes	Yes	Yes	Yes
Ginestra 2023	Can't tell	Yes	Can't tell	Yes	Yes
Hare 2021	Can't tell	Yes	Can't tell	Yes	Yes
Harewood 2009	Can't tell	Yes	Yes	No	Yes
Hsiang 2019	Yes	Yes	Yes	Yes	Yes
Hueston 1996	Can't tell	Yes	Yes	Yes	Yes
Hughes 2019	Yes	Yes	Yes	Can't tell	Yes
Hunt 2021	Yes	Yes	Can't tell	Yes	Yes
Kaneshiro 2010	Can't tell	Yes	Yes	Yes	Yes
Keswani 2016	Can't tell	Yes	Can't tell	Yes	Yes
Kim 2018	Can't tell	Yes	Yes	Yes	Yes
Kolla 2023	Can't tell	Yes	Can't tell	Yes	Yes
Kushnir 2019	Can't tell	Yes	Yes	No	Yes
Lee 2017	Can't tell	Yes	Can't tell	Yes	Yes
Leffler 2012	Can't tell	Yes	Yes	Yes	Yes
Linder 2014	Can't tell	Yes	Can't tell	Yes	Yes
Long 2011	Can't tell	Yes	Yes	Yes	Yes
Lu 2023	Can't tell	Yes	Can't tell	Yes	Yes
Lurix 2012	Can't tell	Yes	Yes	Yes	Yes
Marcondes 2018	Yes	Yes	Yes	Yes	Yes
McClelland 2013	Can't tell	Yes	No	No	Yes
Mitchell 2006	Can't tell	Yes	Can't tell	No	Yes

Munson 2011	Can't tell	Yes	Can't tell	No	Yes
Neprash 2019	Can't tell	Yes	Yes	Yes	Yes
Oakes 2021	Can't tell	Yes	Yes	Yes	Yes
Paeck 2013	Yes	Yes	No	Yes	Yes
Persson 2019	Yes	Yes	Can't tell	Yes	Yes
Philpot 2018	Can't tell	Yes	Can't tell	Yes	Yes
Pu 2019	Yes	Yes	Yes	Yes	Yes
Rawshani 2017	Can't tell	Yes	No	Yes	Yes
Sapci 2022	Can't tell	Yes	Can't tell	No	Yes
Soukup 2018	Can't tell	Yes	Can't tell	Yes	Yes
Soukup 2020a	Yes	Yes	Yes	Yes	Yes
Soukup 2020b	Yes	Yes	Yes	Yes	Yes
Stecker 2015	Can't tell	Yes	No	Yes	Yes
Taylor-Phillips 2015	No	Yes	Can't tell	Yes	Yes
Thurtle 2014	Can't tell	Yes	Can't tell	Yes	Yes
Trinh 2021	No	No	Can't tell	Yes	Yes
Vosshenrich 2021	Can't tell	Yes	No	No	Yes
Wright 2006	Can't tell	Yes	No	Yes	Yes
Xu 2018	Can't tell	Yes	Yes	No	Yes
Zhang 2023a	Can't tell	Yes	Can't tell	No	Yes
Zhang 2023b	Can't tell	Yes	Can't tell	No	Yes
Zheng 2020	Can't tell	Yes	Can't tell	Yes	Yes
Quantitative Descriptive Studies					

Methodological quality criteria for quantitative descriptive studies	Is the sampling strategy relevant to address the research question?	Is the sample representative of the target population?	Are the measurements appropriate?	Is the risk of nonresponse bias low?	Is the statistical analysis appropriate to answer the research question?
Pignatiello 2022	No	No	Yes	Can't tell	Yes

Opinions, narrative overviews, and editorials (using the JBI Opinion Checklist)							
Study (Author & Year)	**Article Type**	Is the **source** of the opinion clearly identified?	Does the source of opinion have **standing in the field of expertise**?	Are the **interests of the relevant population** the central focus of the opinion?	Is the stated position the result of an **analytical process**, and is there **logic in the opinion expressed**?	Is there **reference** to the **extant literature**?	Is any **incongruence** with the literature/sources logically **defended**?
Dubash 2020	Conceptual Overview/Editorial	Yes	Yes	Yes	Yes	Yes	Unclear
Hatami 2022	Editorial	Yes	Unclear	Unclear	Yes	Unclear	No
Lee 2013	Narrative Overview	Yes	Unclear	Yes	Yes	Yes	Yes
Lee 2021	Book Chapter	Yes	Yes	Yes	Yes	Yes	Yes
Moorhouse 2020	Opinion Piece	Yes	Yes	Yes	Yes	No	No
Nasa 2023	Editorial	Yes	Unclear	Yes	Yes	No	No
Oakes 2021	Opinion Piece	Yes	Unclear	Yes	Yes	Yes	Yes
Schweitzer 2023	Personal viewpoint	Yes	Yes	Yes	Yes	Yes	Yes

Literature Reviews (using the JBI Review Checklist)	
Checklist items / Study (Author & Year)	Soukup 2019
Is the review question clearly and explicitly stated?	Yes
Were the inclusion criteria appropriate for the review question?	Yes
Was the search strategy appropriate?	Yes
Were the sources and resources used to search for studies adequate?	Yes
Were the criteria for appraising studies appropriate?	Not applicable
Was critical appraisal conducted by two or more reviewers independently?	No
Were there methods to minimize errors in data extraction?	Unclear
Were the methods used to combine studies appropriate?	Yes
Was the likelihood of publication bias assessed?	No
Were recommendations for policy and/or practice supported by the reported data?	Not applicable
Were the specific directives for new research appropriate?	Yes

Quantitative and qualitative studies were rated using the MMAT which discourages the calculation of overall scores (Hong et al., 2018). The qualitative study fulfilled all five methodological quality criteria (Cygler et al., 2021). One of three randomised controlled trials also fulfilled all five quality criteria (Taylor-Phillips et al., 2016), while the other two received 'can't tell' and 'no' ratings for one to three out of five criteria. Most included quantitative studies fell into the 'Quantitative non-randomised' study design category. As a result of our inclusion criteria, all studies scored 'yes' for the fifth quality criterion ('During the study period, is the intervention administered (or exposure occurred) as intended?'). A majority were rated as 'can't tell' for the first criterion ('Are the participants representative of the target population?'), reflecting that information on the healthcare professionals included in most studies was limited and therefore prevented the assessment. Interestingly the patient populations were generally better described, however, our focus on healthcare professionals' decision-making, made health professionals the population of interest. No studies received more than two 'no' ratings out of five methodological quality criteria.

Opinion pieces, narrative overviews, and editorials were all rated using the JBI Opinion Checklist. Four out of the eight articles received 'yes' ratings for all, or all but one of the criteria. One opinion piece and one editorial received two ‘no’ ratings as reference to extant literature was missing and any potential incongruence with the literature was not addressed. Another editorial received three ‘unclear’ ratings and one ‘no’ rating indicating a low level of quality overall.

The included literature review was rated using the JBI Review Checklist. As the review was not a systematic review, multiple criteria received ‘not applicable’ or ‘no’ ratings.

Discussion

Main findings

We systematically reviewed the literature and identified 82 studies assessing decision fatigue in healthcare professionals. This is the first systematic review to examine empirical evidence relating exclusively to the decision fatigue phenomenon. By incorporating diverse and disconnected strands of literature including those explicitly investigating ‘decision fatigue’ and those using comparable designs but using alternative language, we provide a comprehensive overview and synthesis of previously isolated evidence. Complementing Pignatiello et al.’s (2020) conceptual review of the decision fatigue phenomenon, our systematic review provides a robust overview of empirical evidence on the presence and consequences of decision fatigue in healthcare.

The phenomenon is clearly recognised by healthcare professionals themselves, with 10 articles presenting expert discussions of the impact of decision fatigue on medical decision-making. This recognition is primarily concentrated in the emergency medicine domain. A clear lack of qualitative research was identified with only one qualitative study meeting the inclusion criteria, highlighting a notable gap in the literature. While there are many qualitative studies on related topics, such as general fatigue, decision-making under stress, and compassion fatigue, qualitative work directly examining decision fatigue (i.e., changes in decision-making that occur as decision burden accumulates over time) is scarce. This limits the field’s understanding of the subjective experiences of decision fatigue in healthcare. The 72 studies that quantitatively assessed decision fatigue effects included 130 individual tests of the decision fatigue hypothesis. Statistically significant support for the decision fatigue hypothesis was found in 45% of cases, but the evidence was mixed. A considerable proportion of cases (32%) reported results not in support of the decision fatigue hypothesis, and 23% of cases had results that were inconclusive, warranting further investigation. No clear differences were apparent between studies that did and did not, find evidence of decision fatigue in terms of study design, healthcare setting, professional group or decision type, although studies finding significant decision fatigue effects had higher levels of methodological rigour on average as assessed with the MMAT (Mixed Methods Appraisal Tool).

Overall, the evidence presents a mixed picture. For example, individual studies identified that the ordering of specific diagnostic tests, imaging tests, cancer screenings tests, and CT scans of the head and abdomen all reduce over time (or over a series of patients, depending on the study) in line with decision fatigue, but others found that the ordering of tests (in general) does not change.

Similarly, detection rates (indicating diagnostic sensitivity) declined in some studies, with examples including adenoma detection rates (ADR), polyp yields in colonoscopies, endoscopy biopsy rates, and lesion detection rates, but not in all studies. For some endoscopy readings, there were no changes detected in polyp/lesion/adenoma detection rates, and for mammography reading decisions, two studies showed a gradual improvement in performance throughout the day contrary to the decision fatigue hypothesis. More detailed insights into how reading sessions are structured, how much experience readers have, and how they are compensated (rewarded) for their work may shed light on why findings differ.

This inconsistency is potentially attributable to the limited way in which decision fatigue has been operationalised in the literature. For example, without a clear understanding of what constitutes the ‘less effortful’ decision in different scenarios, it is impossible to make firm predictions. Ordering a test could be effortless in one context (e.g., where tests can be ordered at the click of a button), and effortful in another (e.g., where a detailed justification must be provided). If this is the case, it is likely important to consider each type of test ordering as a unique context, and avoid generalising across tests in general. Similarly, some outcomes used to test decision fatigue hypotheses in the literature are decision-related rather than decisions per se. For example, detection rates in scan reading arguably reflect perceptual vigilance more than decision-making.

Declines in safety-related behaviours were also observed in this review, including that nurses were slower to react to physiological monitor alarms with each successive hour of their shifts, and hand hygiene compliance dropped over the course of a shift. However, other studies reported increased odds of recall rates for diagnostic procedures and increasingly conservative triaging decisions over the shift, both of which are arguably safer, indicating that 'less effortful' decisions are not necessarily less risky decisions. Again, this is likely to be context-specific. For example, where decisions can be directly linked to an individual, there may be more of a tendency to avoid risk with increasing decision fatigue and 'off-load' the patient to another professional to defer responsibility; however, when the path of responsibility is less clear, for example where a professional is one of many who could respond, the opposite effect might take place.

In multidisciplinary oncology team meetings (aka tumour boards, generally consisting of histopathologists, radiologists, surgeons, specialist cancer nurses, and oncologists), the quality of presented patient information, the quality of disciplinary contributions and the quality of team interactions decreased, suggesting that effortful, proactive elements of interpersonal communication are negatively affected by decision fatigue. Some team interaction aspects however did not change, including arguably more reactive aspects: giving answers to questions, negative reactions and showing antagonism.

In terms of prescribing medications, statin prescribing and vaccination rates decreased and opioid and antibiotic prescribing increased as the day advanced, indicating decreases in preventive and increases in reactive treatment decisions. One potential explanation is that providing reactive medications to patients who are ill or in pain is likely to be less effortful than explaining to the patient why they are not receiving treatment, and how to self-manage their condition instead, whilst preventive actions are more likely to require additional effort to explain and justify to the patient in the absence of any current symptoms. However, again the effects were not ubiquitous. The timing of appointments did not affect the prescribing of NSAIDs or antihypertensives.

In other domains, later case start times were associated with higher probabilities of intraoperative red blood cell transfusions and delayed extubations in surgery and anaesthesia, and patients meeting surgeons later in their shifts were less likely to be scheduled for operations. However, elsewhere, the probabilities of transfusions for plasma, cryoprecipitate, or platelets did not change. The rate of primary treatment delivery errors and the ordering of referrals to physical therapy also showed no significant variation throughout the day. The probability of a patient being administered a transfusion was found to change significantly with later case start times but the directionality of change was not discussed and therefore no conclusion can be drawn as to whether this was in support of the decision fatigue hypothesis. This again highlights the need for clear and theoretically justified a priori predictions about expected effects and the relative effort associated with different options in each specific context under study. Specific contexts, such as those involving highly structured clinical pathways, appear less affected by decision fatigue, possibly due to the standardised nature of the processes involved. This raises the important point that decision fatigue hypotheses should only be tested in situations where it is possible in principle for a clinician to deviate from a standard decision, i.e., where there is some decision latitude.

While decision fatigue is often associated with negative outcomes, the effects of a shift towards less effortful decision making is likely to be highly context-dependent. In some cases for example, decision fatigue may lead to more conservative decision-making, such as opting for safer treatment options, increasing diagnostic recall rates, or deferring complex decisions to avoid potential errors (Allan et al., 2019). This can, in turn, result in safer patient care and potentially better patient outcomes in certain cases but may also contribute to inefficient resource allocation across the healthcare system, such as unnecessary testing, overuse of precautionary measures, or increased delays in care. Conversely, in contexts where inaction or reduced effort is associated with suboptimal patient outcomes, such as declining to order necessary tests or failing to adhere to safety protocols, decision fatigue may have detrimental effects. This underscores the importance of considering decision fatigue within the context of the specific cognitive and structural demands of different medical contexts. Future research should aim to describe the conditions under which decision fatigue results in harmful, neutral, or even beneficial outcomes for patients, the clinicians who make the decisions, and the overall healthcare system.

Measurement of decision fatigue

Studies included in this review operationalised and measured decision fatigue in many different ways, making a meta-analysis impossible. The lack of standardisation across studies hinders comprehensive inferences about the general magnitude of decision fatigue effects or even insight into when and how such an effect can be expected to develop. It also raises questions about whether the measurement methods influence the detection of any effects that are present.

A primary goal in studying decision fatigue is to observe repeated decision-making, ideally with clear start – and endpoints and without breaks, to detect changes in *cumulative* decision-making over defined periods of time. Decision fatigue is inherently a within-person phenomenon, developing within people over time as they become decision fatigued, so it is vital to be able to track how far into the decision-making period people are. However, most existing research is carried out with routinely collected healthcare data, where decisions can often not be tied to individual healthcare professionals. Consequently, researchers must analyse population averages without knowing the precise start and end points of individuals' decision-making periods. While shift patterns can be approximated and, in some cases, estimates can be made about break timings during shifts, the variability in healthcare professionals' schedules makes this approximation approach suboptimal, leading to potential misclassification of decisions (into early or later periods of repeated decision-making) and a reduced ability to detect any effect that is present. Breaks have also been shown to 'reset' the decision fatigue effect (Allan et al., 2019; Linder et al., 2014). Without controlling for breaks, studies may combine data from individuals who made decisions continuously with those who took breaks, again leading to a potential underestimation of, or inability to detect decision fatigue effects. Future studies should aim to ensure that decisions can be linked to individual healthcare professionals, or that clear information on shift start and end points (and break timings when possible) is available so that time into shift can be averaged in a valid way. Statistical analyses/ models should aim to account for the clustering of observations within individuals. It is also important to be able to order decisions chronologically within the period of interest. Both the number of patients seen and time into a shift have been used in the reviewed literature as measures of a healthcare professional's decision-making burden at a certain point. Studies that group decisions by broader time windows, such as morning versus afternoon, were often identified in our search, but were excluded (unless the time periods were short) as they do not adequately capture changes in decision-making over time. Where possible, future studies should use the order number of patients seen / problems dealt with, or, if not feasible, time into the shift for ordering decisions chronologically. For time-based measures, clustering decisions by one-hour windows across the work period may be the most practical approach as this has been implemented most widely in existing studies. However, time-based proxies fail to account for variations in number of patients seen during a certain time period, preventing valid comparisons between contexts with frequent short consultations like primary care, and those with fewer, longer patient encounters like surgery.

Future studies should also begin to assess potential real-time confounders such as tiredness, hunger and stress. These are not routinely measured in existing observational studies but would be expected to co-occur with decision fatigue, potentially leading to a misattribution of any observed effects. Structured observations, ecological momentary assessment (EMA) studies, and contextual inquiries could help in gaining an understanding of such confounders.

While a unidimensional self-report measure of decision fatigue exists in the form of the decisional fatigue scale (DFS), it is not widely used in the literature reviewed and has only been validated among surrogate decision-makers of the critically ill (Hickman et al., 2018), and among nurses (Pignatiello et al., 2022). The 9-item scale (modified from an original 10-item scale) was designed to capture 'the influence of dysregulation in emotion regulation (emotional distress), cognitive processing (mental exhaustion), and impulsive decision-making (behaviours) as perceived by the respondent across 1 week.' (Hickman et al., 2018). This differs from the predominant definition of decision fatigue in the literature as a tendency towards making less effortful decisions, as the cumulative burden of effortful decision-making increases. It also focuses on a week-long period (or 24-hour period according to Pignatiello et al. (2022)), which conflicts with the prevailing view in the field that decision fatigue effects accumulate within a relatively short defined period of effortful decision-making, for example during a clinic session or over a single work shift.

At the time of writing, the DFS remains the only self-report measure that attempts to assess subjective aspects of decision fatigue. Its use has been limited to date, and it has a narrow validation base, underscoring the need for the development of new measurement tools that better reflect the current conceptualisation of decision fatigue as a dynamic, within-person process that unfolds over periods of sustained decision-making. Without such tools, researchers remain limited in their ability to capture the phenomenon as it occurs in real-world healthcare contexts.

Theoretical underpinnings

A key finding of the current review is that there are substantial weaknesses in how decision fatigue has been theoretically conceptualised in the existing literature. Our review found evidence of decision fatigue effects on medical decision-making across a wide range of healthcare contexts, emphasising the practical implications of this phenomenon. However, a significant portion of the literature lacks either a clear definition or a robust theoretical grounding. Many articles do not provide any theoretical basis for their hypotheses and/or findings, and those that do often keep this very brief. This is problematic because theory is essential for advancing the field; it guides the formulation of hypotheses, the design of studies, and the interpretation of findings.

Despite being the focus of this review, the term 'decision fatigue' was only used in 35% of the included papers, indicating a lack of consistent terminology within the literature. Some papers used broader terms like 'time of day effect', while others adopted narrower, context-specific terminology such as 'operator fatigue'. This variation is understandable, given that the literature originates from diverse clinical fields with different disciplinary perspectives and norms. However, inconsistent and imprecise terminology hinders clarity and precision, complicates data synthesis, impacts reproducibility, and delays the cumulative advancement of knowledge. To address these issues, the research community should aim to standardise terminology related to decision fatigue.

When defining the concept, there is considerable overlap in the definitions used, suggesting that researchers have broadly comparable views of the phenomenon and are aiming to test similar questions. However, nuanced differences indicate that there is not yet a sufficiently precise or widely accepted definition. The lack of a standardised definition remains a major limitation in this field, hindering theoretical advancement and comparability across studies. To address this gap, we propose that decision fatigue should be defined as 'a tendency towards making less effortful decisions as the cumulative burden of effortful decision-making increases', as this reflects the prevailing view in the literature reviewed. By offering this definition, we aim to provide greater conceptual clarity and facilitate more consistent operationalisation in future research. This proposed definition does not make theoretical claims as to why the effect arises but this should be incorporated once a clearer understanding of the underlying mechanisms has developed.

There is limited use of theoretical frameworks in the decision fatigue literature captured in this review. Only a minority of studies provide a theoretical grounding to their research, and those that do, most commonly reference the Strength/Resource Model of Self-Control (commonly referred to as 'ego depletion') (Baumeister, 2003; Baumeister et al., 2007); conceptualising decision fatigue as an effort-related depletion in the mental resources required to make decisions. While effort-related resource depletion offers a plausible explanation of the decision fatigue effect, it is unlikely to be a feasible explanation. The Strength Model of Self Control has received considerable critique in recent years, most notably when high-profile and large-scale registered replications failed to demonstrate any measurable ego depletion effect under controlled conditions, casting doubt on the robustness of the theory (Hagger et al., 2016). Critics have pointed out methodological flaws in the original studies supporting ego depletion (Carter et al., 2015), there is ambiguity regarding what exactly constitutes the 'resource' that gets depleted (Inzlicht & Schmeichel, 2012) and depletion effects can be reduced with incentives (Muraven & Slessareva, 2003) which indicates a motivational, rather than depletion based explanation for the effect. These and further critiques make it unlikely that ego depletion can be considered a feasible explanation of the decision fatigue effect.

Alternative theoretical frameworks mentioned in the literature reviewed, which could explain decision fatigue, include process-focused (Inzlicht et al., 2014; Inzlicht & Schmeichel, 2012) opportunity cost models (Thaler & Sunstein, 2003). Process models propose that expending effort over time (where this involves an element of effortful control or regulation) results in changes in attention and motivation that orient people away from their current task towards alternatives that are more rewarding or gratifying. Opportunity cost models propose that as people expend effort, they begin to unconsciously weigh up the costs of continued task persistence against the potential benefits of other things they could be doing with the same mental effort. Both suggest a motivational rather than resource-related explanation for changes in effort over time, essentially proposing, in contrast to the ego depletion model, that reductions in performance don't occur because people are unable to exert further effort; rather, they occur because people become less motivated to do so. A motivational explanation aligns with classic perspectives on work, such as Hull's Law of Less Work (1943) which recognises that where alternative courses of action require different amounts of effort, people will learn over time to choose the less effortful action. While Hull's Law refers to physical work, the principle can be readily extended to cognitive work as humans are known to be 'cognitive misers' (Taylor, 1981) who seek to minimise cognitive effort in a drive for economy. A bias that functions to minimise the cognitive effort that is expended over time would be adaptive in the sense that it would steer people towards efficient solutions (Botvinick, 2007) and retain valuable cognitive resources for new or unexpected tasks. Finally, decision fatigue has also been linked to standalone concepts like default bias from more general theories of choice architecture (e.g., Thaler and Sunstein (2003)). Default bias suggests that individuals have a general preference for pre-set options, as the default tends to require less effort or action to select. Consequently, it is proposed that this tendency will increase in situations where decision makers are fatigued or busy.

Although the patterns observed in the decision fatigue literature align broadly with predictions from all of these different theoretical frameworks, it is important to discriminate between them because each theory proposes different underlying mechanisms and, in turn, different potential interventions for decision fatigue. For instance, the Strength/Resource Model suggests that healthcare professionals have limited control over depletion; much as a battery loses its charge over time, under this model, only rest would replenish the depleted resources. This implies that decision fatigue interventions should focus on the provision of recovery periods at strategic points in the working day. Frameworks which emphasise motivational shifts or opportunity costs, in contrast, would suggest that decision fatigue could be managed by enhancing motivation (e.g., through the provision of incentives or rewards) or by limiting access to alternative activities during periods of task focus (e.g., by blocking access to easier or more gratifying activities while a challenging task is being completed). Choice architecture theories would suggest a further distinct option, that decision fatigue interventions should ignore the individual and instead focus on modifying the choice environment to align the desired action with the default action (e.g., by altering the default options in electronic health records to the most beneficial or necessary choice).

The current decision fatigue literature is largely atheoretical, in that explanatory theories and potential mechanisms are suggested in only a small minority of studies. Future studies must begin to systematically test theory-based predictions more robustly, to distinguish between resource, motivational and bias-related explanations of the effect. As outlined above, appropriate interventions can only be designed once the underlying mechanisms of the effect are identified. In addition, valid hypotheses can only be generated and measures developed when researchers have a fuller understanding of when and why decision fatigue would be expected to occur.

Strengths and limitations of the present review

A key strength of our systematic review is its contribution to the literature. To our knowledge, this is the first systematic review of decision fatigue as a standalone concept, and the results therefore offer a valuable overview of the current evidence base. Another strength is our implementation of a highly comprehensive search strategy. Recognising the emergent nature of the field and the diverse terminologies used across different disciplines, we used an inclusive search strategy to capture relevant literature from a range of disciplinary areas. This enabled us to investigate decision fatigue across a multitude of contexts, encompassing various healthcare settings (across primary and secondary care), professional groups (such as nurses, primary care clinicians, and medical specialists), geographical areas with differing healthcare systems, and types of medical decisions. We identified only one study conducted in a psychiatric setting, and none from fields such as audiology, dentistry, ophthalmology, pharmacy, or psychology. This suggests that decision fatigue remains an under-researched phenomenon in these healthcare contexts. We highlight this as a gap in the literature and suggest that future research could usefully explore how decision fatigue manifests in these areas, particularly given their distinct clinical demands and decision-making environments.

Most studies focused on medical decisions related to the evaluation or interpretation of test results (most often related to endoscopies), decisions about medications (such as prescribing of opioids and statins), or therapeutic procedures (including the transfusion of blood products, influenza vaccinations, or the activation of hospital stroke alert systems). By classifying decisions systematically using the DICTUM categories, we were able to identify gaps that future research should address to fully understand the breadth and depth of decision fatigue's effects in clinical practice. A final strength lies in our adherence to gold standard systematic review methodology, including pre-registering the review, pre-specifying our inclusion and exclusion criteria, conducting screening and data extraction partly in duplicate to reduce errors, and assessing the quality of included studies. This rigorous approach enhances the reliability and validity of our findings. Despite our robust search strategy, our review also has limitations. Notably, we restricted our inclusion to English-language publications due to resource limitations and lack of funding to translate non-English studies. This may have limited the generalisability of our findings and likely contributed to the predominance of studies from the US and UK in our review. Additionally, like many systematic reviews, we are dependent on the quality and reporting of the original studies included. Variations in study design, sample sizes, and methodologies can introduce heterogeneity, which can affect the comparability of results and the overall conclusions drawn. Finally, publication bias is a potential limitation. Studies with positive or significant results are more likely to be published, while those with null or negative findings might remain unpublished. This could skew our review's conclusions if such studies were not identified during our search process.

Practical implications

The findings of this review have several important practical implications for healthcare systems, institutions, and individual practitioners. Given the evidence that decision fatigue can impact clinical decision-making, it is crucial to consider strategies with the potential to mitigate its effects and promote high-quality, efficient healthcare delivery. One potential approach is structuring work schedules and breaks to minimise the accumulation of decision fatigue. Introducing mandatory rest periods, ensuring adequate shift lengths, and scheduling complex decision-making tasks earlier in shifts may help maintain cognitive performance. Additionally, implementing strategic workload distribution, such as rotating decision-intensive tasks among team members, could help balance cognitive effort throughout a shift. Decision-support tools and automated prompts within the electronic health record may offer another avenue for mitigating decision fatigue. Automated decision aids can assist clinicians in making evidence-based choices by reducing cognitive load and minimising reliance on effortful deliberation in fatigued states. For example, reminders to offer preventive care or alerts that ask for a justification when commonly overprescribed drugs are selected may help to counteract the tendency toward less effortful decisions. Training and awareness programmes could also play a role in addressing decision fatigue. Educating healthcare professionals about the phenomenon, its potential consequences, and strategies to manage it could empower them to recognise and mitigate its effects in their own practice. Similar strategies have been used successfully to reduce the impact of other cognitive biases on medical decision-making (Croskerry et al., 2013). Encouraging reflective practice and fostering a culture of shared decision-making, where complex cases are discussed collaboratively, may also help distribute decision-making burdens more effectively.

On a broader scale, policy interventions and workflow redesign should consider how system-level changes can reduce the risk of decision fatigue. This might include optimising appointment scheduling, ensuring adequate staffing levels, and integrating redundancy checks for critical decisions to safeguard against fatigue-related errors. Finally, while decision fatigue is often associated with negative outcomes, its context-dependent nature suggests that interventions need to be tailored to specific medical settings and decision types. Future research should focus on identifying high-risk scenarios where decision fatigue is most detrimental and experimentally testing targeted strategies to enhance decision resilience while maintaining efficiency in resource use. By implementing these practical strategies, healthcare organisations can improve decision-making quality, enhance patient safety, and optimise resource allocation, ultimately contributing to more effective and sustainable healthcare delivery.

While the strategies proposed above have potential, their implementation is likely to be limited by logistical, financial, and institutional barriers. Restructuring shifts or integrating decision-support tools often requires significant resource and system-wide coordination, which may not be feasible in already under-funded and understaffed clinical settings. Similarly, it may be impractical to implement regular lengthy breaks in services where patient numbers and demand for appointments are high. However, smaller-scale interventions may offer viable alternatives. For example, micro-breaks during surgeries have been shown to improve the mental focus of surgeons without extending procedure time (Hallbeck et al., 2017), and simple electronic health record-based tools, such as prescribing alerts or preventive screening and care reminders (Onders et al., 2014), may be able to reduce cognitive load by integrating reminders and prompts into the existing electronic health record infrastructure (Kim et al., 2018). These approaches highlight that it may be possible to use modest, well-integrated changes to help mitigate decision fatigue in routine clinical practice.

Conclusion

Our systematic review marks the first comprehensive analysis of decision fatigue among healthcare professionals, revealing significant fatigue-related changes in medical decision-making across diverse medical decision-making contexts. Despite methodological limitations in the current body of literature, decision fatigue is shown to impact various decisions, including diagnostic test orders, medication prescribing, and therapeutic procedures in ways that have the potential to impact healthcare quality and efficiency. However, the evidence remains fragmented and inconsistent, necessitating further research across underexplored decision types and varied healthcare settings. Inconsistencies in concept definition and measurement across the evidence base hinder comprehensive inferences, and explanatory theoretical frameworks are underdeveloped, meaning that many studies lack theoretical underpinnings. To address this gap, we propose defining decision fatigue as 'a tendency towards making less effortful decisions as the cumulative burden of effortful decision-making increases.' Future research should explore more precise ways of measuring decision burden, explore the impact of potential confounders and develop and systematically test predictions from different plausible theoretical models to identify the mechanisms that underlie the decision fatigue effect. In addition to refining theoretical understanding, more qualitative research is needed to explore the experiential and contextual aspects of decision fatigue in healthcare professionals. The current literature is heavily reliant on quantitative methodologies, and the lack of qualitative studies limits insights into how decision fatigue is perceived, managed, and experienced in real-world clinical settings. Addressing this gap will provide a more nuanced understanding of decision fatigue and inform the development of targeted interventions. Advancing our understanding of the effect will enable appropriate interventions such as scheduled breaks, prompts within electronic health records, decision support tools, or strategic workload management to be selected and tested with a view to mitigating any negative impacts of decision fatigue on healthcare delivery.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by a University of Aberdeen Institute of Applied Health Sciences PhD Studentship awarded to MM.

ORCID

Mona Maier http://orcid.org/0000-0001-8270-261X

References

JBI Manual for Evidence Synthesis. (2020). JBI. https://doi.org/10.46658/JBIMES-20-01

Addis, D. R., Moore, B. A., Garner, C. R., Fernando, R. J., Kim, S. M., & Russell, G. B. (2020). Case start time affects intraoperative transfusion rates in adult cardiac surgery: A single-center retrospective analysis. *Journal of Cardiothoracic and Vascular Anesthesia*, *34*(3), 632–639. https://doi.org/10.1053/j.jvca.2019.10.044

Agarwal, I., Joseph, J. W., & Sanchez, L. D. (2022). Time on shift in the emergency department and decision to prescribe opioids to patients without chronic opioid use. *Clinical and Experimental Emergency Medicine*, *9*(2), 108–113. https://doi.org/10.15441/ceem.22.212

Allan, J. L., Johnston, D. W., Powell, D. J. H., Farquharson, B., Jones, M. C., Leckie, G., & Johnston, M. (2019). Clinical decisions and time since rest break: An analysis of decision fatigue in nurses. *Health Psychology*, *38*, 318–324. https://doi.org/10.1037/hea0000725

Almadi, M. A., Sewitch, M., Barkun, A. N., Martel, M., & Joseph, L. (2015). Adenoma detection rates decline with increasing procedural hours in an endoscopist's workload. *Canadian Journal of Gastroenterology and Hepatology*, *29*, 304–308. https://doi.org/10.1155/2015/789038

Anastasian, Z. H., Gaudet, J. G., Levitt, L. C., Mergeche, J. L., Heyer, E. J., & Berman, M. F. (2014). Factors that correlate with the decision to delay extubation following multi-level prone spine surgery. *Journal of Neurosurgical Anesthesiology*, *26*(2), 167–171. https://doi.org/10.1097/ANA.0000000000000028

Backmann, H. A., Larsen, M., Danielsen, A. S., & Hofvind, S. (2021a). Does it matter for the radiologists' performance whether they read short or long batches in organized mammographic screening? *European Radiology*, *31*(12), 9548–9555. https://doi.org/10.1007/s00330-021-08010-9

Backmann, H. A., Larsen, M., Danielsen, A. S., & Hofvind, S. (2021b). Time of day and mammographic reader performance in a population-based breast cancer screening programme. *Journal of Medical Screening*, *28*(3), 295–301. https://doi.org/10.1177/0969141320953206

Baer, T., & Schnall, S. (2021). Quantifying the cost of decision fatigue: Suboptimal risk decisions in finance. *Royal Society Open Science*, *8*(5), 201059. https://doi.org/10.1098/rsos.201059

Baumeister, R. F. (2003). Ego depletion and self-regulation failure: A resource model of self-control. *Alcoholism: Clinical and Experimental Research*, *27*(2), 281–284. https://doi.org/10.1097/01.ALC.0000060879.61384.A4

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science*, *16*(6), 351–355. https://doi.org/10.1111/j.1467-8721.2007.00534.x

Beg, S., Card, T., Sidhu, R., Wronska, E., Ragunath, K., Ching, H.-L., Koulaouzidis, A., Yung, D., Panter, S., Mcalindon, M., Johnson, M., Kurup, A., Shonde, A., San-Juan Acosta, M., Sansone, S., Simmon, E., Thurston, V., Healy, A., Chetcuti Zammit, S., … Tiwari, R. (2021). The impact of reader fatigue on the accuracy of capsule endoscopy interpretation. *Digestive and Liver Disease*, *53*(8), 1028–1033. https://doi.org/10.1016/j.dld.2021.04.024

Bernstein, M. H., Baird, G. L., & Lourenco, A. P. (2022). Digital breast tomosynthesis and digital mammography recall and false-positive rates by time of day and reader experience. *Radiology*, *303*(1), 63–68. https://doi.org/10.1148/radiol.210318

Bersani, K., Fuller, T. E., Garabedian, P., Espares, J., Mlaver, E., Businger, A., Chang, F., Boxer, R. B., Schnock, K. O., Rozenblum, R., Dykes, P. C., Dalal, A. K., Benneyan, J. C., Lehmann, L. S., Gershanik, E. F., Bates, D. W., & Schnipper, J. L. (2020). Use, perceived usability, and barriers to implementation of a patient safety dashboard integrated within a vendor EHR. *Applied Clinical Informatics*, *11*(1), 34–45. https://doi.org/10.1055/s-0039-3402756

Bonafide, C. P., Localio, A. R., Holmes, J. H., Nadkarni, V. M., Stemler, S., MacMurchy, M., Zander, M., Roberts, K. E., Lin, R., & Keren, R. (2017). Video analysis of factors associated with response time to physiologic monitor alarms in a children's hospital. *JAMA Pediatrics*, *171*(6), 524–531. https://doi.org/10.1001/jamapediatrics.2016.5123

Bostock, N. (2016, May 27). *Exclusive: Three quarters of GPs take no breaks during a four-hour clinical session*. GPonline. https://www.gponline.com/article/1396723

Botvinick, M. M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(4), 356–366. https://doi.org/10.3758/CABN.7.4.356

Canfield, C., Udeh, C., Blonsky, H., Hamilton, A. C., & Fertel, B. S. (2020). Limiting the number of open charts does not impact wrong patient order entry in the emergency department. *Journal of the American College of Emergency Physicians Open*, *1*(5), 1071–1077. https://doi.org/10.1002/emp2.12129

Carter, E. C., Kofler, L. M., Forster, D. E., & McCullough, M. E. (2015). A series of meta-analytic tests of the depletion effect: Self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General*, *144*(4), 796–815. https://doi.org/10.1037/xge0000083

Castrén, M., Kurland, L., Liljegard, S., & Djärv, T. (2015). Non-specific complaints in the ambulance; predisposing structural factors. *BMC Emergency Medicine*, *15*(1), 8. https://doi.org/10.1186/s12873-015-0034-5

Chan, M. Y., Cohen, H., & Spiegel, B. M. R. (2009). Fewer polyps detected by colonoscopy as the day progresses at a veteran's administration teaching hospital. *Clinical Gastroenterology and Hepatology*, *7*(11), 1217–1223. https://doi.org/10.1016/j.cgh.2009.07.013

Changolkar, S., Rareshide, C. A. L., Snider, C. K., & Patel, M. S. (2020). Patient, physician, and environmental predictors of influenza vaccination during primary care visits. *Journal of General Internal Medicine*, *35*(2), 611–613. https://doi.org/10.1007/s11606-019-05017-3

Cheng, V. C., Tai, J. W., Ho, S. K., Chan, J. F., Hung, K. N., Ho, P. L., & Yuen, K. Y. (2011). Introduction of an electronic monitoring system for monitoring compliance with moments 1 and 4 of the WHO 'My 5 moments for hand hygiene' methodology. *BMC Infectious Diseases*, *11*(1), 151. https://doi.org/10.1186/1471-2334-11-151

Christensen, E. E., Dietz, G. W., Murry, R. C., & Moore, J. G. (1977). The effect of fatigue on resident performance. *Radiology*, *125*(1), 103–105. https://doi.org/10.1148/125.1.103

Croskerry, P. (2017). Medical decision making. In L. J. Ball & V. A. Thompson (Eds.), *International handbook of thinking and reasoning*. Routledge.

Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 2: Impediments to and strategies for change. *BMJ Quality & Safety*, *22*(Suppl 2), ii65–ii72. https://doi.org/10.1136/bmjqs-2012-001713

Cygler, J., Page, A. V., & Ginsburg, S. (2021). Life on call: Perspectives of junior and senior internal medicine residents. *Academic Medicine : Journal of the Association of American Medical Colleges*, *96*(5), 744–750. https://doi.org/10.1097/ACM.0000000000003803

Dai, H., Milkman, K. L., Hofmann, D. A., & Staats, B. R. (2015). The impact of time at work and time off from work on rule compliance: The case of hand hygiene in health care. *Journal of Applied Psychology*, *100*(3), 846–862. https://doi.org/10.1037/a0038067

Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, *108*(17), 6889–6892. https://doi.org/10.1073/pnas.1018033108

Delaney, K. R., & Fogg, L. (2005). Patient characteristics and setting variables related to use of restraint on four inpatient psychiatric units for youths. *Psychiatric Services*, *56*(2), 186–192. https://doi.org/10.1176/appi.ps.56.2.186;

Department of Health, L. T. C. T. (2012). *Long term conditions compendium of information: Third edition*. GOV.UK. https://www.gov.uk/government/publications/long-term-conditions-compendium-of-information-third-edition

Dexheimer, J. W., Kirkendall, E. S., Kouril, M., Hagedorn, P. A., Minich, T., Duan, L. L., Mahdi, M., Szczesniak, R., & Spooner, S. A. (2017). The effects of medication alerts on prescriber response in a pediatric hospital. *Applied Clinical Informatics*, *8*(2), 491–501. https://doi.org/10.4338/ACI-2016-10-RA-0168

Dominello, M. M., Paximadis, P., Zaki, M., Hammoud, A., Campbell, S., Komajda, M., Dyson, G., Bossenberger, T., & Burmeister, J. (2015). Ten-year trends in safe radiation therapy delivery and results of a radiation therapy quality assurance intervention. *Practical Radiation Oncology*, *5*(6), e665–e671. https://doi.org/10.1016/j.prro.2015.08.004

Dubash, R., Bertenshaw, C., & Ho, J. H. (2020). Decision fatigue in the emergency department. *Emergency Medicine Australasia*, *32*(6), 1059–1061. https://doi.org/10.1111/1742-6723.13670

Edwards, H., Courtney, M., Wilson, J., Monaghan, S., & Walsh, A. (2003). Fever management audit: Australian nurses' antipyretic usage. *Pediatric Nursing*, *29*(1), 31–37.

Ergün-Şahin, B., Güneş, E. D., Kocabıyıkoğlu, A., & Keskin, A. (2022). How does workload affect test ordering behavior of physicians? An empirical investigation. *Production and Operations Management*, *31*(6), 2664–2680. https://doi.org/10.1111/poms.13711

Evans, M. I., Richardson, D. A., Sholl, J. S., & Johnson, B. A. (1984). Cesarean section. Assessment of the convenience factor. *The Journal of Reproductive Medicine*, *29*(9), 670–676.

Faulk, C. E., Cooper, N. R., Staneata, J. A., Bunch, M. P., Galang, E., Fang, X., & Foster, K. J. (2013). Rate of return to acute care hospital based on day and time of rehabilitation admission. *PM&R*, *5*(9), 757–762. https://doi.org/10.1016/j.pmrj.2013.06.002

Fisher, R. F., Croxson, C. H., Ashdown, H. F., & Hobbs, F. R. (2017). GP views on strategies to cope with increasing workload: A qualitative interview study. *British Journal of General Practice*, *67*(655), e148–e156. https://doi.org/10.3399/bjgp17X688861

Freedman, J. S., Harari, D. Y., Bamji, N. D., Bodian, C. A., Kornacki, S., Cohen, L. B., Miller, K. M., & Aisenberg, J. (2011). The detection of premalignant colon polyps during colonoscopy is stable throughout the workday. *Gastrointestinal Endoscopy*, *73*(6), 1197–1206. https://doi.org/10.1016/j.gie.2011.01.019

Ginestra, J. C., Kohn, R., Hubbard, R. A., Auriemma, C. L., Patel, M. S., Anesi, G. L., Kerlin, M. P., & Weissman, G. E. (2023). Association of time of day with delays in antimicrobial initiation among ward patients with hospital-onset sepsis. *Annals of the American Thoracic Society*, *20*(9), 1299–1308. https://doi.org/10.1513/AnnalsATS.202302-160OC

Glöckner, A. (2016). The irrational hungry judge effect revisited: Simulations reveal that the magnitude of the effect is overestimated. *Judgment and Decision Making*, *11*(6), 601–610. https://doi.org/10.1017/S1930297500004812

Graber, M. (2005). Diagnostic errors in medicine: A case of neglect. *The Joint Commission Journal on Quality and Patient Safety*, *31*(2), 106–113. https://doi.org/10.1016/S1553-7250(05)31015-4

Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., Brand, R., Brandt, M. J., Brewer, G., Bruyneel, S., Calvillo, D. P., Campbell, W. K., Cannon, P. R., Carlucci, M., Carruth, N. P., Cheung, T., Crowell, A., De Ridder, D. T. D., Dewitte, S., … Zwienenberg, M. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science*, *11*(4), 546–573. https://doi.org/10.1177/1745691616652873

Hallbeck, M. S., Lowndes, B. R., Bingener, J., Abdelrahman, A. M., Yu, D., Bartley, A., & Park, A. E. (2017). The impact of intraoperative microbreaks with exercises on surgeons: A multi-center cohort study. *Applied Ergonomics*, *60*, 334–341. https://doi.org/10.1016/j.apergo.2016.12.006

Hare, A. J., Adusumalli, S., Park, S., & Patel, M. S. (2021). Assessment of primary care appointment times and appropriate prescribing of statins for at-risk patients. *JAMA Network Open*, *4*(5), e219050. https://doi.org/10.1001/jamanetworkopen.2021.9050

Harewood, G. C., Chrysostomou, K., Himy, N., & Leong, W. L. (2009). Impact of operator fatigue on endoscopy performance: Implications for procedure scheduling. *Digestive Diseases and Sciences*, *54*(8), 1656–1661. https://doi.org/10.1007/s10620-008-0549-7

Hatami, Z., Sarkhani, N., & Nikpeyma, N. (2022). Decision fatigue in nurses in the COVID-19 pandemic: A commentary—nursing open—2021—Hatami—decision fatigue in nurses in the COVID-19 pandemic A commentary.pdf. *Nursing Open*, 4–5.

Hickman, R. L., Jr., Pignatiello, G. A., & Tahir, S. (2018). Evaluation of the decisional fatigue scale among surrogate decision makers of the critically Ill. *Western Journal of Nursing Research*, *40*(2), 191–208. https://doi.org/10.1177/0193945917723828

Hirshleifer, D., Levi, Y., Lourie, B., & Teoh, S. H. (2019). Decision fatigue and heuristic analyst forecasts. *Journal of Financial Economics*, *133*(1), 83–98. https://doi.org/10.1016/j.jfineco.2019.01.005

Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., Vedel, I., & Pluye, P.. (2018). The mixed methods appraisal tool (MMAT) version 2018 for information professionals and researchers. *Education for Information*, *34*(4), 285–291. https://doi.org/10.3233/EFI-180221

Hsiang, E. Y., Mehta, S. J., Small, D. S., Rareshide, C. A. L., Snider, C. K., Day, S. C., & Patel, M. S. (2019). Association of primary care clinic appointment time with clinician ordering and patient completion of breast and colorectal cancer screening. *JAMA Network Open*, *2*(5), e193403. https://doi.org/10.1001/jamanetworkopen.2019.3403

Hueston, W. J., McClaflin, R. R., Claire, E., & Falls, S. (1996). Variations in cesarean delivery for fetal distress. *The Journal o f Family Practice*, *43*(5), 461–467.

Hughes, J., Lysikowski, J., Acharya, R., Phelps, E., & Kandil, E. (2020). A multi-year analysis of decision fatigue in opioid prescribing. *Journal of General Internal Medicine*, *35*(4), 1337–1339. https://doi.org/10.1007/s11606-019-05217-x

Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory* (p. x, 422). Appleton-Century.

Hunt, T. C., Ambrose, J. P., Haaland, B., Kawamoto, K., Dechet, C. B., Lowrance, W. T., Hanson, H. A., & O'Neil, B. B. (2021). Decision fatigue in low-value prostate cancer screening. *Cancer*, *127*(18), 3343–3353. https://doi.org/10.1002/cncr.33644

Inzlicht, M., & Schmeichel, B. J. (2012). What is ego depletion? Toward a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science*, *7*(5), 450–463. https://doi.org/10.1177/1745691612454134

Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences*, *18*(3), 127–133. https://doi.org/10.1016/j.tics.2013.12.009

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, *93*(5), 1449–1475. https://doi.org/10.1257/000282803322655392

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*(4), 637–671. https://doi.org/10.3758/BF03196323

Kaneshiro, M., Ho, A., Chan, M., Cohen, H., & Spiegel, B. M. (2010). Colonoscopy yields fewer polyps as the day progresses despite using social influence theory to reverse the trend. *Gastrointestinal Endoscopy*, *72*(6), 1233–1240. https://doi.org/10.1016/j.gie.2010.08.034

Keswani, R. N., Gawron, A. J., Cooper, A., & Liss, D. T. (2016). Procedure delays and time of day are not associated with reductions in quality of screening colonoscopies. *Clinical Gastroenterology and Hepatology*, *14*(5), 723–728.e2. https://doi.org/10.1016/j.cgh.2015.10.023

Kim, R. H., Day, S. C., Small, D. S., Snider, C. K., Rareshide, C. A. L., & Patel, M. S. (2018). Variations in influenza vaccination by clinic appointment time and an active choice intervention in the electronic health record to increase influenza vaccination. *JAMA Network Open*, *1*(5), 1–10. https://doi.org/10.1001/jamanetworkopen.2018.1770

Kolla, L., Chen, J., & Parikh, R. B. (2023). Time of clinic appointment and serious illness communication in oncology. *Cancer Control: Journal of the Moffitt Cancer Center*, *30*, 10732748231170488. https://doi.org/10.1177/10732748231170488

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *The Behavioral and Brain Sciences*, *36*(6), 661–679. https://doi.org/10.1017/S0140525X12003196

Kushnir, A., Bleznak, J. L., Saslow, J. G., & Stahl, G. (2020). Nurses' finnegan scoring of newborns with neonatal abstinence syndrome not affected by time or Day of the week. *American Journal of Perinatology*, *37*(2), 224–230. https://doi.org/10.1055/s-0039-1698458

Lee, X. Q. (2021). Decision fatigue effect. In M. Raz & P. Pouryahya (Eds.), *Decision making in emergency medicine* (pp. 103–110). Springer. https://doi.org/10.1007/978-981-16-0143-9_17

Lee, A., Jensen, C. D., Marks, A. R., Zhao, W. K., Doubeni, C. A., Zauber, A. G., Quinn, V. P., Levin, T. R., & Corley, D. A. (2017). Endoscopist fatigue estimates and colonoscopic adenoma detection in a large community-based setting. *Gastrointestinal Endoscopy*, *85*(3), 601–610.e2. https://doi.org/10.1016/j.gie.2016.09.033

Lee, C. S., Nagy, P. G., Weaver, S. J., & Newman-Toker, D. E. (2013). Cognitive and system factors contributing to diagnostic errors in radiology. *American Journal of Roentgenology*, *201*(3), 611–617. https://doi.org/10.2214/AJR.12.10375

Leffler, D. A., Kheraj, R., Bhansali, A., Yamanaka, H., Neeman, N., Sheth, S., Sawhney, M., Lamont, J. T., & Aronson, M. D. (2012). Adenoma detection rates vary minimally with time of day and case rank: A prospective study of 2139 first screening colonoscopies. *Gastrointestinal Endoscopy*, *75*(3), 554–560. https://doi.org/10.1016/j.gie.2011.11.021

Linder, J. A., Doctor, J. N., Friedberg, M. W., Reyes Nieva, H., Birks, C., Meeker, D., & Fox, C. R. (2014). Time of day and the decision to prescribe antibiotics. *JAMA Internal Medicine*, *174*(12), 2029–2031. https://doi.org/10.1001/jamainternmed.2014.5225

Long, M. D., Martin, C., Sandler, R. S., Herfarth, H. H., Shaheen, N. J., & Dellon, E. S. (2011). Reduced polyp detection as endoscopy shift progresses experience with screening colonoscopy at a tertiary-care hospital. *Journal of Clinical Gastroenterology*, *45*(3), 253–258. https://doi.org/10.1097/MCG.0b013e3181fd2998

Lu, Z., Zhang, L., Yao, L., Gong, D., Wu, L., Xia, M., Zhang, J., Zhou, W., Huang, X., He, C., Wu, H., Zhang, C., Li, X., & Yu, H. (2023). Assessment of the role of artificial intelligence in the association between time of day and colonoscopy quality. *JAMA Network Open*, *6*(1), e2253840. https://doi.org/10.1001/jamanetworkopen.2022.53840

Lurix, E., Hernandez, A. V., Thoma, M., & Castro, F. (2012). Adenoma detection rate is not influenced by full-day blocks, time, or modified queue position. *Gastrointestinal Endoscopy*, *75*(4), 827–834. https://doi.org/10.1016/j.gie.2011.12.008

Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, *1*(1), 6–21. https://doi.org/10.1080/17470214808416738

Maier, M., Allan, J. L., Powell, D., & Murchie, P. (2021). *Systematic review of the effects of decision fatigue in healthcare professionals on medical decision making*. PROSPERO 2024. https://www.crd.york.ac.uk/PROSPERO/view/CRD42021260081.

Marcondes, F. O., Gourevitch, R. A., Schoen, R. E., Crockett, S. D., Morris, M., & Mehrotra, A. (2018). Adenoma detection rate falls at the end of the day in a large multi-site sample. *Digestive Diseases and Sciences*, *63*(4), 856–859. https://doi.org/10.1007/s10620-018-4947-1

McClelland, L. E., Switzer, F. S., III, & Pilcher, J. J. (2013). Changes in nurses' decision making during a 12-h day shift. *Occupational Medicine*, *63*(1), 60–65. https://doi.org/10.1093/occmed/kqs189

Mehlmann-Wicks, J. (2022). *Safe working in general practice*. The British Medical Association Is the Trade Union and Professional Body for Doctors in the UK. https://www.bma.org.uk/advice-and-support/gp-practices/managing-workload/safe-working-in-general-practice

Mitchell, H., Hocking, J., & Saville, M. (2006). Temporal characteristics of laboratory screening errors in cervical cytology. *Acta Cytologica*, *50*, 492–498. https://doi.org/10.1159/000326002

Moorhouse, A. (2020). Decision fatigue: Less is more when making choices with patients. *British Journal of General Practice*, *70*(697), 399. https://doi.org/10.3399/bjgp20X711989

Morgan, R. L., Whaley, P., Thayer, K. A., & Schünemann, H. J. (2018). Identifying the PECO: A framework for formulating good questions to explore the association of environmental and other exposures with health outcomes. *Environment International*, *121*(Pt 1), 1027–1031. https://doi.org/10.1016/j.envint.2018.07.015

Munson, G. W., Harewood, G. C., & Francis, D. L. (2011). Time of day variation in polyp detection rate for colonoscopies performed on a 3-hour shift schedule. *Gastrointestinal Endoscopy*, *73*(3), 467–475. https://doi.org/10.1016/j.gie.2010.07.025

Muraven, M., & Slessareva, E. (2003). Mechanism of self-control failure: Motivation and limited resources. *Personality and Social Psychology Bulletin*, *29*(7), 894–906. https://doi.org/10.1177/0146167203029007008

Nasa, P., & Majeed, N. A. (2023). Decision fatigue among emergency physicians: Reality or myth. *Indian Journal of Critical Care Medicine: Peer-Reviewed, Official Publication of Indian Society of Critical Care Medicine*, *27*(9), 609–610. https://doi.org/10.5005/jp-journals-10071-24526

Neprash, H. T., & Barnett, M. L. (2019). Association of primary care clinic appointment time with opioid prescribing. *JAMA Network Open*, *2*(8), 1–11. https://doi.org/10.1001/jamanetworkopen.2019.10373

Oakes, A. H., Adusumalli, S., Snider, C. K., Rareshide, C. A. L., & Patel, M. S. (2021). Variation in cardiologist statin prescribing by clinic appointment time. *Journal of the American College of Cardiology*, *77*(5), 661–662. https://doi.org/10.1016/j.jacc.2020.11.057

Oakes, A. H., & Patel, M. S. (2021). Time to address disparities in care by appointment time. *Healthcare*, *9*(1), 100507. https://doi.org/10.1016/j.hjdsi.2020.100507

Ofstad, E. H., Frich, J. C., Schei, E., Frankel, R. M., & Gulbrandsen, P. (2016). What is a medical decision? A taxonomy based on physician statements in hospital encounters: A qualitative study. *BMJ Open*, *6*(2), e010098. https://doi.org/10.1136/bmjopen-2015-010098

Ofstad, E. H., Frich, J. C., Schei, E., Frankel, R. M., Šaltytė Benth, J., & Gulbrandsen, P. (2018). Clinical decisions presented to patients in hospital encounters: A cross-sectional study using a novel taxonomy. *BMJ Open*, *8*(1), e018042. https://doi.org/10.1136/bmjopen-2017-018042

Onders, R., Spillane, J., Reilley, B., & Leston, J. (2014). Use of electronic clinical reminders to increase preventive screenings in a primary care setting: Blueprint from a successful process in Kodiak, Alaska. *Journal of Primary Care & Community Health*, *5*(1), 50–54. https://doi.org/10.1177/2150131913496116

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Paeck, K. H., Heo, W. J., Park, D. I., Kim, Y.-H., Lee, S.-H., Lee, C. K., Eun, C. S., & Han, D. S. (2013). Colonoscopy scheduling influences adenoma and polyp detection rates. *Hepato-Gastroenterology*, *60*(127), 1647–1652.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

Persson, E., Barrafrem, K., Meunier, A., & Tinghög, G. (2019). The effect of decision fatigue on surgeons' clinical decision making. *Health Economics*, *28*(10), 1194–1203. https://doi.org/10.1002/hec.3933

Philpot, L. M., Khokhar, B. A., Roellinger, D. L., Ramar, P., & Ebbert, J. O. (2018). Time of day is associated with opioid prescribing for low back pain in primary care. *Journal of General Internal Medicine*, *33*(11), 1828–1830. https://doi.org/10.1007/s11606-018-4521-8

Pignatiello, G. A., Martin, R. J., & Hickman, R. L. (2020). Decision fatigue: A conceptual analysis. *Journal of Health Psychology*, *25*(1), 123–135. https://doi.org/10.1177/1359105318763510

Pignatiello, G. A., Tsivitse, E., O'Brien, J., Kraus, N., & HickmanJr.R. L. (2022). Decision fatigue among clinical nurses during the COVID-19 pandemic. *Journal of Clinical Nursing*, *31*(7–8), 869–877. https://doi.org/10.1111/jocn.15939

Pu, L., Lu, K., Ovenden, A., Rana, K., Singh, G., Krishnamurthi, S., Edwards, S., Wilson, B., Nakamura, M., Yamamura, T., Ruszkiewicz, A., Hirooka, Y., Burt, A. D., & Singh, R. (2019). Effect of time of day and specialty on polyp detection rates in Australia. *Journal of Gastroenterology and Hepatology*, *34*(5), 899–906. https://doi.org/10.1111/jgh.14566

Rawshani, A., Rawshani, N., Gelang, C., Andersson, J. O., Larsson, A., Bång, A., Herlitz, J., & Gellerstedt, M. (2017). Emergency medical dispatch priority in chest pain patients due to life threatening conditions: A cohort study examining circadian variations and impact of the education. *International Journal of Cardiology*, *236*, 43–48. https://doi.org/10.1016/j.ijcard.2017.02.047

Sapci, I., Jia, X., Wu, J., Gorgun, E., Liska, D., Church, J., Steele, S. R., & Valente, M. A. (2022). Does time of day matter for colonoscopy quality? A review of over 13,000 screening colonoscopies in a colorectal surgery unit. *The American Journal of Surgery*, *223*(3), 500–502. https://doi.org/10.1016/j.amjsurg.2021.11.032

Saposnik, G., Maurino, J., Sempere, A. P., Terzaghi, M. A., Ruff, C. C., Mamdani, M., Tobler, P. N., & Montalban, X. (2017). Overcoming therapeutic inertia in multiple sclerosis care: A pilot randomized trial applying the traffic light system in medical education. *Frontiers in Neurology*, *8*(430), 1–12. https://doi.org/10.3389/fneur.2017.00430

Schweitzer, D. R., Baumeister, R., Laakso, E.-L., & Ting, J. (2023). Self-control, limited willpower and decision fatigue in healthcare settings. *Internal Medicine Journal*, *53*(6), 1076–1080. https://doi.org/10.1111/imj.16121

See, J. E., Howe, S. R., Warm, S., & Dember, W. N. (1995). *Meta-analysis of the sensitivity decrement in vigilance*.

Soukup, T., Gandamihardja, T. A. K., McInerney, S., Green, J. S. A., & Sevdalis, N. (2019a). Do multidisciplinary cancer care teams suffer decision-making fatigue: An observational, longitudinal team improvement study. *BMJ Open*, *9*(5), e027303. https://doi.org/10.1136/bmjopen-2018-027303

Soukup, T., Lamb, B. W., Morbi, A., Shah, N. J., Bali, A., Asher, V., Gandamihardja, T., Giordano, P., Darzi, A., Sa Green, J., & Sevdalis, N. (2020a). A multicentre cross-sectional observational study of cancer multidisciplinary teams: Analysis of team decision making. *Cancer Medicine*, *9*(19), 7083–7099. https://doi.org/10.1002/cam4.3366

Soukup, T., Lamb, B. W., Shah, N. J., Morbi, A., Bali, A., Asher, V., Gandamihardja, T., Giordano, P., Darzi, A., Green, J. S. A., & Sevdalis, N. (2020b). Relationships between communication, time pressure, workload, task complexity, logistical issues and group composition in transdisciplinary teams: A prospective observational study across 822 cancer cases. *Frontiers in Communication*, *5*(583294), 1–12. https://www.frontiersin.org/articles/10.3389fcomm.2020.583294

Soukup, T., Lamb, B. W., Weigl, M., Green, J. S. A., & Sevdalis, N. (2019b). An integrated literature review of time-on-task effects with a pragmatic framework for understanding and improving decision-making in multidisciplinary oncology team meetings. *Frontiers in Psychology*, *10*(1245), 1–21. https://doi.org/10.1111/1467-9450.00192

Stecker, M. M., Michel, K., Antaky, K., Wolin, A., & Koyfman, F. (2015). Characteristics of the stroke alert process in a general hospital. *Surgical Neurology International*, *6*(1), 5. https://doi.org/10.4103/2152-7806.149387

Stewart, A. F., Ferriero, D. M., Josephson, S. A., Lowenstein, D. H., Messing, R. O., Oksenberg, J. R., Johnston, S. C., & Hauser, S. L. (2012). Fighting decision fatigue. *Annals of Neurology*, *71*(1), A5–A15. https://doi.org/10.1002/ana.23531

Taylor-Phillips, S., Elze, M. C., Krupinski, E. A., Dennick, K., Gale, A. G., Clarke, A., & Mello-Thoms, C. (2015). Retrospective review of the drop in observer detection performance over time in lesion-enriched experimental studies. *Journal of Digital Imaging*, *28*(1), 32–40. https://doi.org/10.1007/s10278-014-9717-9

Taylor-Phillips, S., Wallis, M. G., Jenkinson, D., Adekanmbi, V., Parsons, H., Dunn, J., Stallard, N., Szczepura, A., Gates, S., Kearins, O., Duncan, A., Hudson, S., & Clarke, A. (2016). Effect of using the same vs different order for second readings of screening mammograms on rates of breast cancer detection A randomized clinical trial. *JAMA-Journal of the American Medical Association*, *315*(18), 1956–1965. https://doi.org/10.1001/jama.2016.5257

Taylor, S. E. (1981). The interface of cognitive and social psychology. *Cognition, Social Behavior, and the Environment*, *1*, 189–211.

Thaler, R. H., & Sunstein, C. R. (2003). Libertarian paternalism. *American Economic Review*, *93*(2), 175–179. https://doi.org/10.1257/000282803321947001

Thurtle, D., Pullinger, M., Tsigarides, J., McIntosh, I., Steytler, C., & Beales, I. (2014). Colonoscopic polyp detection rate is stable throughout the workday including evening colonoscopy sessions. *F1000Research*, *3*(107). https://doi.org/10.12688/f1000research.4045.1

Trinh, P., Hoover, D. R., & Sonnenberg, F. A. (2021). Time-of-day changes in physician clinical decision making: A retrospective study. *PLoS One*, *16*(9), e0257500. https://doi.org/10.1371/journal.pone.0257500

Vosshenrich, J., Brantner, P., Cyriac, J., Boll, D. T., Merkle, E. M., & Heye, T. (2021). Quantifying radiology resident fatigue: Analysis of preliminary reports. *Radiology*, *298*(3), 632–639. https://doi.org/10.1148/radiol.2021203486

Weinshall-Margel, K., & Shapard, J. (2011). Overlooked factors in the analysis of parole decisions. *Proceedings of the National Academy of Sciences*, *108*(42), E833–E833. https://doi.org/10.1073/pnas.1110910108

Wright, M. C., Phillips-Bute, B., Mark, J. B., Stafford-Smith, M., Grichnik, K. P., Andregg, B. C., & Taekman, J. M. (2006). Time of day effects on the incidence of anesthetic adverse events. *BMJ Quality & Safety*, *15*(4), 258–263. https://doi.org/10.1136/qshc.2005.017566

Xu, Y., Chen, K., Xu, L., Yuan, X., Wu, Y., & Chen, P. (2018). Diagnostic yield is not influenced by the timing of screening endoscopy: Morning versus afternoon. *Scandinavian Journal of Gastroenterology*, *53*(3), 365–369. https://doi.org/10.1080/00365521.2018.1433230

Zhang, Z., Chen, X., Cao, T., Ning, Y., Wang, H., Wang, F., Zhao, Q., & Fang, J. (2023a). Polyps are detected more often in early colonoscopies. *Scandinavian Journal of Gastroenterology*, *58*(9), 1085–1090. https://doi.org/10.1080/00365521.2023.2202293

Zhang, Z., Chen, X., Wang, H., Nie, H., Wang, F., Zhao, Q., & Fang, J. (2023b). Esophagogastroduodenoscopy outcomes variated by the time of the Day: A single-center experience. *Journal of Clinical Medicine*, *12*(3), 863–873. https://doi.org/10.3390/jcm12030863

Zheng, B., Kwok, E., Taljaard, M., Nemnom, M. J., & Stiell, I. (2020). Decision fatigue in the emergency department: How does emergency physician decision making change over an eight-hour shift? *American Journal of Emergency Medicine*, *38*(12), 2506–2510. https://doi.org/10.1016/j.ajem.2019.12.020