

# UC Davis

## UC Davis Previously Published Works

### Title

The relationship between physician burnout and quality of healthcare in terms of safety and acceptability: a systematic review.

### Permalink

<https://escholarship.org/uc/item/05n2f21n>

### Journal

BMJ open, 7(6)

### ISSN

2044-6055

### Authors

Dewa, Carolyn S  
Loong, Desmond  
Bonato, Sarah  
et al.

### Publication Date

2017-06-01

### DOI

10.1136/bmjopen-2016-015141

Peer reviewed

# BMJ Open The relationship between physician burnout and quality of healthcare in terms of safety and acceptability: a systematic review

Carolyn S Dewa,<sup>1,2</sup> Desmond Loong,<sup>2</sup> Sarah Bonato,<sup>3</sup> Lucy Trojanowski<sup>2</sup>

**To cite:** Dewa CS, Loong D, Bonato S, *et al*. The relationship between physician burnout and quality of healthcare in terms of safety and acceptability: a systematic review. *BMJ Open* 2017;**7**:e015141. doi:10.1136/bmjopen-2016-015141

► Prepublication history and additional material are available. To view these files please visit the journal online (<http://dx.doi.org/10.1136/10.1136/bmjopen-2016-015141>).

Received 11 November 2016

Revised 24 March 2017

Accepted 21 April 2017



CrossMark

<sup>1</sup>Department of Psychiatry and Behavioral Sciences, University of California, Davis, California, USA

<sup>2</sup>Centre for Research on Employment and Workplace Health, Centre for Addiction and Mental Health, Toronto, Canada

<sup>3</sup>Library Services, Centre for Addiction and Mental Health, Toronto, Canada

## Correspondence to

Dr Carolyn S Dewa; [csdewa@ucdavis.edu](mailto:csdewa@ucdavis.edu)

## ABSTRACT

**Objectives** This study reviews the current state of the published peer-reviewed literature related to physician burnout and two quality of care dimensions. The purpose of this systematic literature review is to address the question, ‘How does physician burnout affect the quality of healthcare related to the dimensions of acceptability and safety?’

**Design** Using a multiphase screening process, this systematic literature review is based on publically available peer-reviewed studies published between 2002 and 2017. Six electronic databases were searched: (1) *MEDLINE Current*, (2) *MEDLINE In-process*, (3) *MEDLINE Epub Ahead of Print*, (4) *PsycINFO*, (5) *Embase* and (6) *Web of Science*.

**Setting** Physicians practicing in civilian settings.

**Participants** Practicing physicians who have completed training.

**Primary and secondary outcome measures** Quality of healthcare related to acceptability (ie, patient satisfaction, physician communication and physician attitudes) and safety (ie, minimising risks or harm to patients).

**Results** 4114 unique citations were identified. Of these, 12 articles were included in the review. Two studies were rated as having high risk of bias and 10 as having moderate risk. Four studies were conducted in North America, four in Europe, one in the Middle East and three in East Asia. Results of this systematic literature review suggest there is moderate evidence that burnout is associated with safety-related quality of care. Because of the variability in the way patient acceptability-related quality of care was measured and the inconsistency in study findings, the evidence supporting the relationship between burnout and patient acceptability-related quality of care is less strong.

**Conclusions** The focus on direct care-related quality highlights additional ways that physician burnout affects the healthcare system. These studies can help to inform decisions about how to improve patient care by addressing physician burnout. Continued work looking at the relationship between dimensions of acceptability-related quality of care measures and burnout is needed to advance the field.

Reports from around the world indicate that about one-third to one-half of physicians experience at least one dimension of burnout.<sup>1–5</sup> Burnout has been conceptualised as a syndrome consisting of three dimensions:

## Strengths and limitations of this study

- Few studies have examined the current state of knowledge about the relationship between physician burnout and the patient safety and acceptability dimensions of quality of care.
- This systematic literature review employed a broad search of six electronic databases: (1) *MEDLINE Current*, (2) *MEDLINE In-process*, (3) *MEDLINE Epub Ahead of Print*, (4) *PsycINFO*, (5) *Embase* and (6) *Web of Science*. A manual search was also conducted. In total, 4114 unique citations were identified and reviewed by three reviewers in pairs.
- We used a comprehensive search strategy that follows the recommended best practices of incorporating adjacency commands and synonyms for keywords.
- One of the limitations of the search strategy employed in this systematic review is its focus on English-language publications.
- Another potential limitation of the search strategy is the focus on published peer-reviewed articles. In doing so, our results may be subject to publication bias.

emotional exhaustion (EE), depersonalisation (DP) and low personal accomplishment (PA).<sup>6</sup> Maslach *et al*<sup>7</sup> define EE as referring to ‘feelings of being overextended and depleted of one’s emotional and physical resources’. DP is also referred to as cynicism and defined as ‘a negative, callous, or excessively detached response to various aspects’.<sup>7</sup> PA is also referred to as professional efficacy and ‘it refers to feelings of incompetence and a lack of achievement and productivity at work’.<sup>7</sup> Burnout has been observed to affect personal well-being through low job satisfaction<sup>8–10</sup> and decreased mental health.<sup>11</sup>

Because physicians play an integral role in the healthcare system, the effects of physician burnout are not limited to the physicians experiencing it. Rather, physician burnout potentially impacts the entire healthcare

system. For example, a recent systematic literature review reported a negative relationship between burnout and productivity (ie, early retirement, work cutback and quitting).<sup>12</sup> The impact of productivity loss related to burnout could lead to fewer available healthcare resources that, in turn, can result in healthcare service waitlists. One estimate of the costs of physician work cutback and early retirement related to burnout suggests it totals to at least \$C213 million in patient services losses.<sup>8</sup>

This raises another question about physicians who continue to practice despite experiencing burnout. Does burnout affect their practice? There is evidence that physician burnout is also related to decreased quality of patient care.<sup>5</sup> The WHO<sup>13</sup> and the Institute of Medicine (IOM)<sup>14</sup> suggest that there are six dimensions for quality of healthcare: effectiveness, efficiency, accessibility, equitability, acceptability and safety.

The purpose of this systematic literature review is to address the question, 'How does physician burnout affect the quality of healthcare related to the dimensions of acceptability and safety?' In this review, we focus on two dimensions of quality: acceptability (ie, patient satisfaction, perceived quality of care and communication) and safety (ie, minimising risks or harm to patients). We chose these two dimensions because they reflect the quality of patient–physician interactions.<sup>15</sup> That is, if a clinician's well-being is compromised, their patient interactions may also be negatively affected.<sup>16</sup> In contrast, effectiveness, efficiency, accessibility and equitability reflect the systems (ie, infrastructure, information technology and payment policies) in which practice is conducted.<sup>14</sup>

## BACKGROUND

There has been growing interest in the relationship between healthcare professional well-being and quality of patient care. Although the WHO<sup>13</sup> and IOM<sup>14</sup> identify six dimensions of quality of healthcare, attention has focused on the dimension of patient safety. Recently, there have been four published reviews that focus on the relationship between healthcare professional well-being and patient safety.<sup>17–20</sup> For example, Hall *et al*<sup>18</sup> consider healthcare staff well-being and Salyers *et al*<sup>20</sup> examine staff burnout as opposed to specifically examining physician burnout as our review does. de Jong *et al*<sup>17</sup> examine common mental disorders as opposed to burnout. Williams and Skinner<sup>19</sup> look at physician satisfaction rather than burnout. Each of these published reviews answers questions that are different from the one addressed in our review. Because they seek to answer different questions, they employ search strategies and inclusion/exclusion criteria that are different from those used in our review. Consequently, they include different articles. For example, Hall *et al*'s<sup>18</sup> review does not include nine articles that are included in our systematic review. Among these, there are six articles related to acceptability and three articles related to patient safety that were not included in Hall *et al*'s<sup>18</sup> review. In comparison to de Jong *et al*'s review,<sup>17</sup> our review has

six articles on acceptability and five on patient safety that are unique to our systematic review. None of the articles included in our review were included in Williams and Skinner's review.<sup>19</sup> Compared with the papers included in Salyers *et al*'s<sup>20</sup> review, there are four papers related to physician burnout and safety that are unique to our review and two focused on acceptability that are unique to our review. Thus, our review includes papers that have not been considered together to look at quality of care related to physician interactions with patients and the impact of burnout on physicians.

In addition, none of the published reviews considers the quality of care dimension of acceptability for physicians who have completed training. Yet, along with patient safety, this dimension reflects the quality of interactions between providers and patients. The physician–patient interactions are one of the fundamental interactions in healthcare.<sup>15 19</sup> Furthermore, the IOM<sup>14</sup> asserts that the rise in chronic illnesses necessitates quality interactions to enhance the collaboration between the physician and patient. Quality of physician–patient interactions is reflected in communication, perceived quality of care and patient satisfaction.<sup>14 15</sup> It is the physician–patient interaction that supports the collaboration that will lead to better patient outcomes.<sup>15</sup>

Wallace *et al*<sup>16</sup> assert that physician well-being could be used as a quality indicator. The argument could be strengthened by also understanding how well-being is associated with the physician–patient interaction-related quality dimensions of safety and acceptability. In particular, burnout could be a focus because it reflects well-being and there are standardised measures to identify it. Furthermore, it is a facet of well-being that can be influenced by organisational factors and is under the influence of the healthcare system.<sup>16 21 22</sup> Thus, this systematic review of the literature extends our knowledge about the dimensions of quality of care that reflect physician interactions with patients and a dimension of well-being that is affected by the work environment.

## METHODS

A systematic review of the literature was reported following the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* guidelines (online supplementary file 1: PRISMA checklist).<sup>23</sup> Ethics board review was not sought because this review relied solely on publicly available sources of information.

## Information sources

Six databases were searched: (1) *MEDLINE Current* (index of biomedical research and clinical sciences journal articles); (2) *MEDLINE In-Process* (index of biomedical research and clinical sciences journal articles awaiting to be indexed into *MEDLINE Current*); (3) *MEDLINE Epub Ahead of Print* (index of articles that appear on publisher websites in advance of the journal release); (4) *PsycINFO* (an index of journal articles, books, chapters and dissertations in psychology, social

sciences, behavioural sciences and health sciences); (5) *Embase* (index of biomedical research and abstracts from biomedical, drug and medical device conferences); and (6) *Web of Science* (index of journal articles, editorially selected books and conference proceedings in life sciences and biomedical research).

### Search strategy

Collaborating with the professional health science librarian (SB) member of this research team, search strategies were developed and tailored for each database following the Peer Review of Electronic Search Strategies guidelines<sup>24</sup> (online supplementary file 2: search terms used in search strategy). Because recommended guidelines were used for this review's search strategies, the search strategy that we used is also a contribution to the literature. As this literature grows, the strategy can be used in future searches on the topic. The searches were conducted in February 2017. The OVID platform was used to search *MEDLINE Current*, *MEDLINE In-Process*, *MEDLINE Epub Ahead of Print*, *PsycINFO* and *Embase*. *Web of Science* was searched using the Thomson Reuters search interface. The search period covered January 2002–February 2017; all searches were limited to English-language journals. The time frame was chosen to represent the current healthcare environments in which physicians are practicing. For example, the year 2002 was the year after the IOM's report<sup>14</sup> on the quality of healthcare that discussed the six dimensions of quality of care. By beginning in 2002, we have allowed for a 1-year lag after publication of this report during which healthcare settings and researchers could have incorporated the IOM's quality of healthcare framework into their work.

Our searches sought to identify articles about practicing physicians regardless of specialty working in civilian settings (ie, non-military settings). In this review, the physician search included: allergists, anaesthesiologists, cardiologists, clinical pharmacologists, clinical toxicologists, dermatologists, doctors, endocrinologists, gastroenterologists, gynaecologists, haematologists, immunologists, medical biochemists, medical geneticists, medical microbiologists, nephrologists, neurologists, neuropathologists, neuroradiologists, occupational physicians, oncologists, ophthalmologists, pathologists, paediatricians, physicians, psychiatrists, radiologists, rheumatologists, surgeons and urologists. The search strategy did not seek to exclude residents and medical students. Rather, a broad search strategy was employed to increase the likelihood that all studies on physician burnout would be found. The reference lists of all accepted full-text articles were hand searched.

### Screening process

Relevant articles were identified using a multiphase screening process that involved reviewer pairs using the inclusion and exclusion criteria for this review. In the first step, titles were screened. Next, abstracts of the articles that remained after the first step were screened. The

final step of the process involved screening the full text of all articles that passed the first and second phases. In the full-text screening, papers for which there was insufficient information in the title and abstract to determine relevancy were also included. Two pairs of reviewers (CSD and LT, CSD and DL) independently completed the multiphase screening process. The inter-rater reliability corrected for chance<sup>25</sup> between CSD and LT, and CSD and DL was  $\kappa=0.96$  and  $\kappa=0.98$ , respectively. Before moving onto each stage, disagreements were discussed until consensus was reached.

For this review, burnout was defined as a syndrome of EE, cynicism (DP) and reduced feelings of PA related to work.<sup>6</sup> Quality of care related to acceptability was identified with measures reflecting physician–patient interactions such as patient satisfaction, perceived quality of care, physician communication with patients and physician attitudes towards patients. In addition, safety was identified by measures that reflected risks or harm to patients such as medical errors.

Study inclusion criteria were:

1. Studies reported quality of care outcomes related to acceptability and/or safety.
2. The sample population was comprised of practicing physicians regardless of specialty who worked in civilian settings. That is, the results were reported such that the practicing physician (as opposed to resident) outcomes were reported separately.
3. Burnout was assessed based on a psychometrically validated measure.
4. Paper reports original research.

Exclusion criteria were:

1. The study sample was comprised only of residents and medical students.
2. The study did not examine the relationship between burnout and one of the two quality of care dimensions.
3. Burnout was not assessed based on a validated measure.
4. The paper was a review article or commentary.

### Risk of bias assessment

All included articles were assessed for risk of bias by both pairs of reviewers (CSD and LT, and CSD and DL). Disagreements between the pairs of reviewers were discussed until consensus was reached.

To assess the risk of bias in observational studies, Sanderson *et al*<sup>26</sup> recommend the use of a transparent checklist that concentrates on the 'few, principal, and potential sources of bias in a study's findings'. They assert that the fundamental domains should include: (1) the appropriate selection of participants, (2) appropriate measurement of variables and (3) appropriate control of confounding. In accordance with their recommendations and the Strengthening of Observational Studies in Epidemiology criteria,<sup>27</sup> a nine-item risk of bias checklist with the following criteria adapted from Lagerveld *et al*<sup>28</sup> was used:

1. Study population is well described to facilitate understanding about the generalisability of the results based on the study sample (eg, age, sex, location of the study, physician specialty and practice location).
2. Data collection methods that address the risk of bias are described.
3. Participation/response rate was at least 50% on average.
4. The psychometric properties of the quality of care outcome measure have been tested.
5. Statistical method was appropriate for the question being answered.
6. Statistical significance of associations were tested and reported.
7. Study controlled for at least one confounder such as sex or age in the analyses.
8. Physician matched with patient.
9. Longitudinal data was used.

Each item was scored '1' if the criterion had been met. Each article could achieve a maximum score of 9. Based on their total score, articles were categorised either as low (8–9 points), moderate (5–7 points) or high risk of bias (1–4 points).

## RESULTS

### Article inclusion and exclusion results

The electronic literature search resulted in the identification of 4114 unique citations (figure 1). Based on the title review, 4020 citations were excluded; this left 94 articles for abstract review. During the abstract review, another 28 citations were excluded; this left 66 articles for full-text review. Reasons for article exclusions at full-text review were: (1) not a relevant outcome (n=10), (2) sample not comprised of physicians/cannot distinguish physicians as a group from other clinicians (n=15), (3) it was not original research (n=20), (4) burnout not measured with a validated instrument (n=1) and (5) not published in a peer-reviewed journal (n=8). After the full-text review, 12 articles remained, and their reference lists were hand searched for relevant studies. The hand search identified six additional citations; all six were excluded at full-text review.

### Risk of bias assessment results

Our assessment indicated 10 of the 12 studies were of moderate risk of bias; two were of high risk of bias. Figure 2 illustrates the limitations of these studies. Two studies comprehensively<sup>5 29</sup> described the study population from which the study sample was drawn. Two studies used longitudinal data.<sup>29 30</sup> Other limitations involved not reporting the response rate<sup>31–34</sup> and not controlling for possible confounding factors in the statistical analyses.<sup>34 35</sup> There was also variability in the use of validated outcome measures; only three studies used validated instruments to measure their outcomes.<sup>31 33 35</sup> All included studies employed appropriate statistical tests. All but one<sup>29</sup> reported the results of the statistical testing

(online supplementary file 3: risk of bias assessment checklist).

### Overview of the studies

Of the 12 studies that met the inclusion criteria (table 1), four were conducted in the USA, two in Germany and one each in Greece, Israel, Japan, China and Taiwan. There was one multinational study based on data from Italy, Spain and Portugal.

### Description of the study populations

Six of the studies focused on hospital-based physicians.<sup>5 30 34–37</sup> Among these studies, two focused on cancer<sup>34</sup> and children's<sup>36</sup> specialty hospitals. In addition, one of these studies recruited surgeons practicing either in general surgery or gynaecological wards.<sup>5</sup> One of these studies<sup>37</sup> also included people practicing as physicians who did not have graduate educations.

The remaining five studies recruited physicians practicing in a variety of settings. Three studies sought physicians in primary healthcare centres<sup>29 31 33</sup>; they included physicians practicing in internal medicine, general practice and family practice. One of the studies<sup>29</sup> that recruited primary care physicians focused on the quality of care only for patients with diabetes and/or hypertension.

Two studies did not specify the setting.<sup>32 38</sup> However, of these two, one focused on surgeons.<sup>38</sup> Finally, one study used four health plans to recruit and contained a mixture of community and hospital physicians,<sup>39</sup> which included physicians specialising in ophthalmology, dermatology, otolaryngology, community-based gynaecology, general surgery and hospital-based cardiology.

### Measuring burnout

In 9 of the 12 studies, burnout was measured using either the 22-item Maslach Burnout Inventory (MBI),<sup>6</sup> translated version of the MBI-GS,<sup>37</sup> translated version of the MBI-HSS<sup>30 31</sup> or selected MBI subscales.<sup>30–38</sup> The complete 22-item MBI measures three dimensions of burnout: EE, DP and PA. It is one of the most widely used measures of burnout in the scientific literature.<sup>40 41</sup> One study<sup>29</sup> used a single-item measure for burnout that correlates with the EE subscale of the MBI.<sup>42</sup>

The two remaining studies used the Copenhagen Burnout Inventory (CBI)<sup>40</sup> and the Shirom-Melamed Burnout Measure (SMBM).<sup>41 43</sup> The CBI is a 19-item scale comprised of three subscales that assess personal burnout, work-related burnout and client-related burnout.<sup>40</sup> It has been shown to be correlated with mental and general health as well as job satisfaction.<sup>40</sup> The SMBM is a 22-item measure with three subscales that assess physical fatigue, EE and cognitive weariness.<sup>41</sup> The psychometric properties of these scales continue to be explored.<sup>41 44 45</sup>

### Measuring quality of care related to acceptability and patient safety

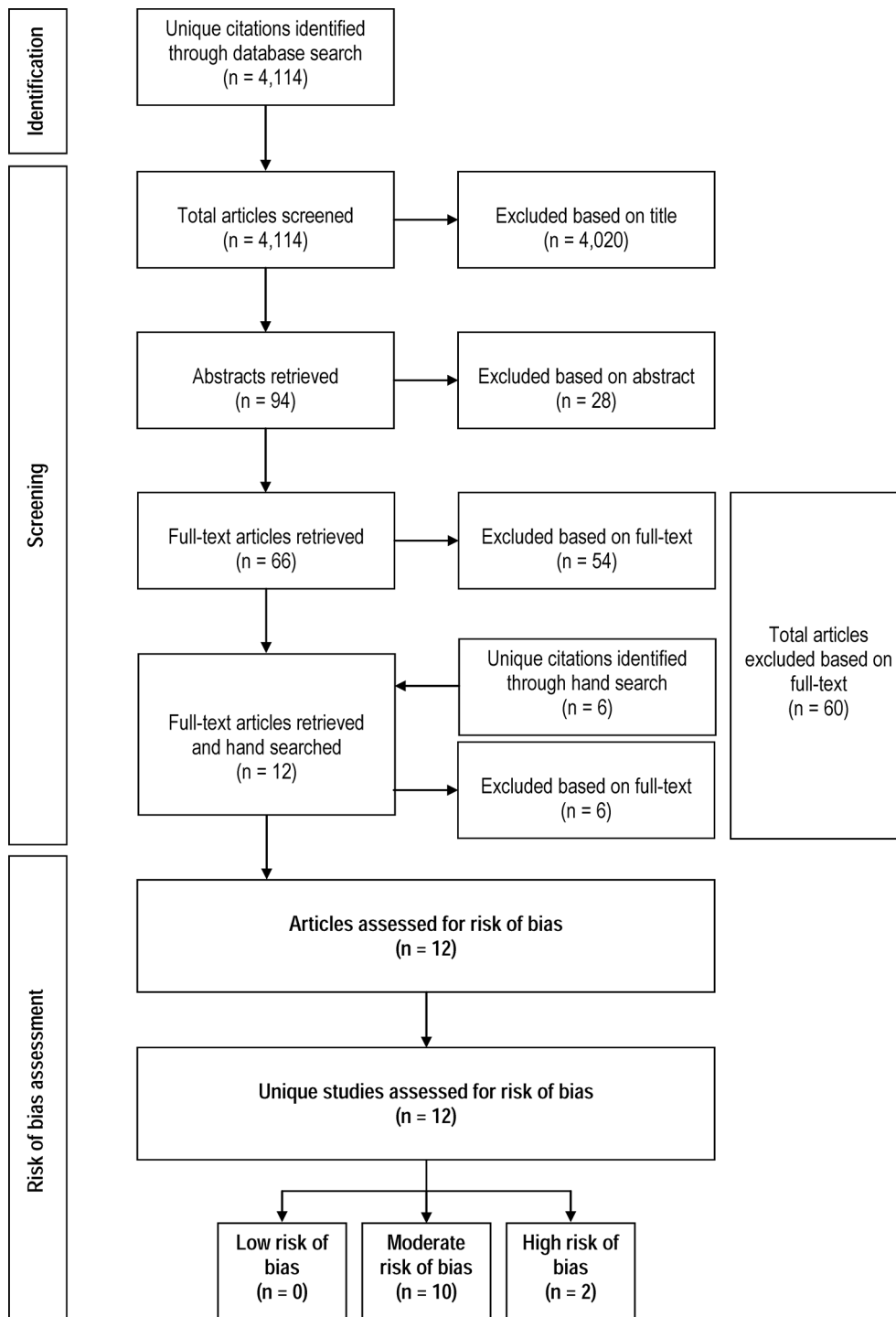
Four types of quality of care measures related to acceptability and safety were used in these studies. In terms of

patient safety, medical errors were measured. Acceptability-related measures included patient satisfaction, perceived general quality of care and physician communication/attitudes.

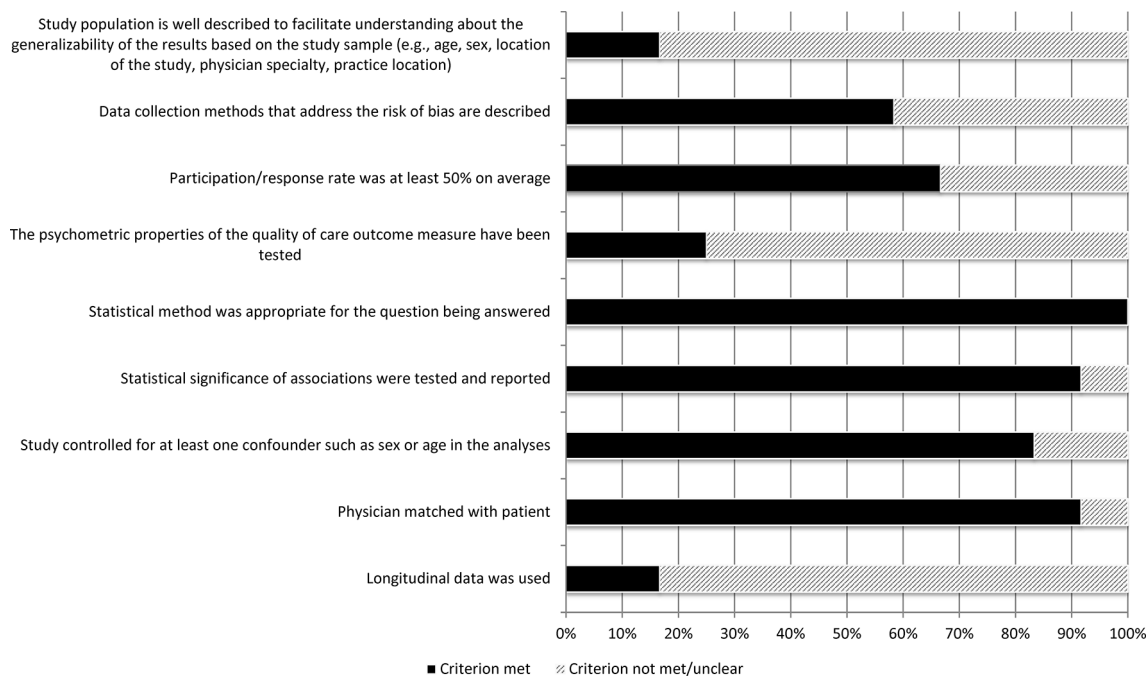
*Patient safety measures: medical errors*

Patient safety was examined with medical errors. This outcome was assessed in five studies.<sup>5 29 30 37 38</sup> Wen *et al*<sup>37</sup> asked respondents whether they had made any medical errors including one that resulted in a patient being

harmed, a medication error, delay in treatment or incomplete or incorrect item being added to the patient record. Hayashino *et al*<sup>30</sup> and Shanafelt *et al*<sup>38</sup> used similar questions about whether the respondent made major medical errors. However, the studies differed in the time frame that the respondent was asked to consider. Hayashino *et al*<sup>30</sup> asked about the past year, while Shanafelt *et al*<sup>38</sup> inquired about the past 3 months. In contrast to these studies, Klein *et al*<sup>5</sup> asked about frequency of diagnostic mistakes and treatment without specifying a time frame. The studies differ in



**Figure 1** Flowchart of accepted/rejected articles.



**Figure 2** Summary of risk of bias across studies.

the types of errors that they asked about (ie, *major* errors rather than *any* errors). In addition, they depend on recall and self-report. Shanafelt *et al*<sup>38</sup> note that studies have used this type of question to gather information about medical errors. However, there are also studies that have found that physicians under-report medical errors.<sup>46</sup> Furthermore, there is evidence that physicians have a limited ability to self-assess their practice patterns.<sup>47</sup>

In addition to questions about frequency of diagnostic mistakes and treatment, Klein *et al*<sup>5</sup> included a questionnaire based on the Canadian Physician Achievement Review to evaluate physician self-perceived quality of psychosocial care, diagnosis/therapy and quality assurance.<sup>48</sup> However, the authors note that additional work regarding its validity is warranted.<sup>5</sup>

There was only one study that did not rely on self-report to gather information about medical errors. Rabatin *et al*<sup>29</sup> used a chart audit to assess medical errors characterised by adherence to guidelines, responsiveness to 'recurrent abnormalities' and missed drug interactions.

#### Acceptability measures: patient satisfaction/perceived quality of care

With regard to acceptability measures, patient satisfaction was assessed in four studies.<sup>31 32 35 39</sup> In two of these studies, the SERVQUAL was used to measure patient satisfaction/quality of care.<sup>32 39</sup> The SERVQUAL was developed to measure service quality along five dimensions: (1) tangibles (ie, physical facilities), (2) reliability (ie, performs dependably and accurately), (3) responsiveness (ie, willingness to help), (4) assurance (ie, ability to inspire trust) and (5) empathy (ie, caring).<sup>49</sup> Halbesleben and Rathert<sup>32</sup> used a healthcare-specific version of the SERVQUAL. The psychometric properties of the scale were examined.<sup>50</sup> However, Asubonteng *et al*<sup>51</sup> have raised

questions about the strength of the scale's psychometric properties.

Shirom and colleagues<sup>39</sup> adapted the SERVQUAL by eliminating seven items and revising the language for physicians to rate their own quality of care using the remaining 15 items. The validity of this modified measure was not examined.

Weigl *et al*<sup>36</sup> looked at physician-perceived quality of care by asking physicians to rate two statements on a five-point scale: 'My workload frequently leads to reduced quality of work' and 'Adverse work conditions frequently lead to a loss of quality.' The authors reference the German version of the MBI as the source for these questions. However, they do not provide information about the psychometric properties of the individual use of these items.

One study<sup>31</sup> used the Consultation Satisfaction Questionnaire (CSQ) scale that was created and validated to assess patient satisfaction with general practitioners.<sup>52</sup> It is comprised of 18 items and measures satisfaction along four dimensions: general satisfaction, professional care, depth of relationship, and perceived time.

Finally, in their study, Weng *et al*<sup>35</sup> used two questions to indicate patient satisfaction, 'I am satisfied with the care provided by my doctor,' and 'I would recommend this doctor to my friends and family.' The first of Weng *et al*'s<sup>35</sup> questions is similar to one of the CSQ's<sup>52</sup> general satisfaction items, 'I am totally satisfied with my visit to the doctor.' However, the use of this single-item has not been validated. A version of the second question has been used to measure satisfaction and was correlated with the EUROPEP patient satisfaction questionnaire.<sup>53</sup>

#### Acceptability measures: communication/attitudes

Two studies focused on physician communication/attitudes.<sup>33 34</sup> Using audiotapes of physician/patient

**Table 1** Study descriptions and reported patient safety and acceptability related quality of care outcomes

Author(s)	Study population	Description of sample	Burnout measure	Quality of care measure	Quality of care outcomes		
					Medical errors (ME)	Patient satisfaction (PS)/quality of care (QoC)	Communication/attitudes
Anagnostopoulos <i>et al.</i> <sup>51</sup> Greece	Physicians working in three large primary healthcare centres. Patients of participating physicians. Patients selected through systematic random sampling – 1:3 consecutive patients. Physician response rate: 85.8% Patient response rate: Not reported	n=30 physicians ≤10 years practicing: 53% Specialities: General practitioners: 63% Pathologists/internists: 23.3% Male: n=17 Female: n=13 >50 years: 43% 26–50 years: 40% n=300 patients Male: 46% Female: 54% Mean age: 54±15 years	Greek translation of the 22-item Maslach Burnout Inventory (MBI)-Human Services Survey	Patient report: Patient satisfaction assessed using 18-item Consultation Satisfaction Questionnaire (CSQ). <sup>52</sup> 5-point Likert scale from 1=‘strongly agree’ to 5=‘strongly disagree’. Satisfaction subscales: (1) general, (2) perceived length of consultation, (3) depth of relationship, and (4) professional care provided Overall satisfaction: sum of all items (max score=90) Scale was translated into Greek using back-translation and pilot testing. English version’s psychometric properties tested. <sup>52,53</sup>	Correlation between MBI dimensions and PS: ▲ EE and PS: $r=-0.64$ , $p<0.01$ ▲ DP and PS: $r=-0.54$ , $p<0.01$ ▲ PA and PS: $r=0.26$ , $p=0.17$ Results of mixed effect model with PS as outcome: ▲ Low EE associated with highest average PS Comparison between moderate and high EE: no significant difference in association with PS	Correlation between MBI dimensions and PS: ▲ DP and PS: $r=-0.16$ , $p<0.05$	
Halbesleben and Rathert <sup>52</sup> USA	Attending physicians of university students who had been hospitalised in past year. Student response rate: Not reported.	n=178 physicians Years practicing: Not reported Specialities: Not reported Male: n=84 Female: n=94 Mean age=46±13 years n=178 patients Male: n=98 Female: n=80 Mean age: 23±5 years	22-item MBI-Human Services Survey modified to apply to patients rather than general care recipients	Patient report: Patient satisfaction assessed using 22-item SERVQUAL. <sup>50</sup> 7-point Likert scale from 1=‘strongly disagree’ to 7=‘strongly agree’. Psychometric properties tested but subsequent study suggested need for further exploration regarding its validity. <sup>51</sup>	Correlation between MBI dimensions and PS: ▲ DP and PS: $r=-0.16$ , $p<0.05$		

Continued



Table 1 Continued

Author(s)	Study population	Description of sample	Burnout measure	Quality of care measure	Quality of care outcomes	
					Medical errors (ME)	Patient satisfaction (PS)/quality of care (QoC)
Hayashino <i>et al</i> <sup>20</sup> Japan	Members of a panel of 6459 hospital-based physicians recruited through hospital lists and scientific meetings. A randomly selected subsample of 1198 were invited to participate. Response rate: 70%	n=836 physicians Years practicing: Not reported Male: 92% Female: 8% 28–39 years: 23% 40–49 years: 47% 50–59 years: 26% 60–81 years: 4%	17-item MBI developed for Japanese healthcare professionals based on the MBI-Human Services Survey Used burnout thresholds: EE: >21 DP: >18 PA: >16	Physician report: Perceived medical errors assessed with questions: 'Are you concerned that you have made any major medical mistakes in the last year?' If 'yes', asked about number of medical errors that concerned respondent. Psychometric properties not tested.	Association between MBI dimensions and any medical error: ▲ Significant differences among tertiles for EE ( $p=0.026$ ) and DP ( $p=0.002$ ) % with ME by burnout dimension tertile: ▲ EE 1st tertile: 27.9% ▲ EE 2nd tertile: 38.2% ▲ EE 3rd tertile: 33.9% ▲ DP 1st tertile: 35.0% ▲ DP 2nd tertile: 27.8% ▲ DP 3rd tertile: 37.2% ▲ No significant differences among tertiles for PA ( $p=0.67$ )	
Klein <i>et al</i> <sup>6</sup> Germany	Physicians in surgery working in >100 beds general hospitals with a general surgical and/or gynaecological ward. Stratified probability sample based on hospital beds. Response rates: Hospital level: 53% Physician level: 36% Physicians in participating hospitals: 65%	n=1311 physicians Mean years practicing: 11 years Male: 60% Female: 40% Mean age=45 ± 8.5 years	Copenhagen Burnout Inventory (CBI). Three scales assessing personal, client and work burnout. This study focused on personal burnout (ie, degree of physical and psychological fatigue and exhaustion).	Physician report: Perceived quality of care assessed using short version of Chirurgisches Qualitätsregister. Created three subscales: (1) psychosocial care, (2) diagnosis/therapy and (3) quality assurance. 5-point Likert scale from 1='very good' to 5='bad'. Two questions about frequency of diagnostic and therapeutic errors: 'I have made mistakes in diagnosis.' and 'I have made mistakes in treatment.' 4-point Likert scale ('never' to 'often'). Psychometric properties not tested for either set of questions.	Adjusted* ORs (95% CI) for probability of suboptimal care and high burnout score: ▲ Psychosocial care=1.58 (1.19, 2.10) ▲ Dx/Tx=1.59 (1.17, 2.16) ▲ Quality assurance=1.45 (1.10, 1.90) *Adjusted for gender, occupational position, job experience	Adjusted* ORs (95% CI) for probability of error and high burnout score: ▲ Diagnostic error: 1.66 (1.26, 2.20) ▲ Therapeutic error: 1.94 (1.39, 2.69) *Adjusted for gender, occupational position, job experience

Continued

**Table 1** Continued

Author(s)	Study population	Description of sample	Burnout measure	Quality of care measure	Quality of care outcomes	
					Medical errors (ME)	Patient satisfaction (PS)/quality of care (QoC)
Rabatin <i>et al</i> <sup>29</sup> USA	Primary care physicians in New York City, Chicago and rural and urban Wisconsin Recruited 1–6 patients/physician with diabetes, hypertension or congestive heart failure. Response rate: Physicians: 59.6% Non-participants did not differ from participants in specialty, age or sex.	n=119 practices n=449 physicians n=1419 patient charts Physician characteristics: Male: n=235 Female: n=187 Mean age: 43±10 years Specialties: Family medicine: 47% General internal medicine: 50% Patient characteristics: Not reported.	Single-item measure: 'Using your own definition of burnout... (a) I have no symptoms of burnout; (b) Occasionally I am under stress... but I don't feel burned out; (c) I am definitely burning out and have one or more symptoms of burnout, such as physical and emotional exhaustion; (d) The symptoms of burnout that I'm experiencing won't go away...; (e) I feel completely burned out and often wonder if I can go on...' The question correlates with the EE dimension of MBI. <sup>42</sup>	Physician report: Chart audit using a standardised template to retrospectively assess over 18 months for guideline adherence, responsiveness to 'recurrent abnormalities' and missed drug interactions. Reliability not reported.	Statistics not reported. No statistically significant differences between physicians with burnout and without.	Communication/attitudes
Ratanawongsa <i>et al</i> <sup>63</sup> USA	Physicians from 15 urban community-based clinics who provided primary care to adult patients enrolled in a randomised controlled trial for hypertensive minority patients. Response rate: Not reported.	n=40 physicians Mean years of practice: 11±7.7 years Male: 47% Female: 53% Mean age: 42±8.7 years Specialties: Internal medicine: 83% Family practice: 15% General practice: 2% n=235 patients Male: 34% Female: 66% Mean age: 59±13.2 years	A 6-item scale derived from the MBI that captures the domains of EE and PA. 5-point Likert scale from 1='strongly agree' to 3='neutral' to 5='strongly agree'. Based on tertiles, burnout scores were categorised as low, average or high.	Physician report: Physicians completed 'brief questionnaires indicating the degree to which they knew the patient, their attitudes toward the patient in general, and their attitudes regarding the visit'. Audiotaped encounters analysed for rapport-building communication behaviours using the Roter Interaction Analysis System. Four types of rapport identified: (1) positive, (2) negative, (3) emotional and (4) social. Reliability and predictive validity tested. <sup>64</sup>	Adjusted* ORs (95% CI) for probability of PS with high versus low burnout: ▶ PS=0.44 (0.18, 1.08), p=0.07 *Adjusted for patient health insurance, visit length, physician gender, physician international medical school graduate (IMG) status, interaction between IMG status and burnout	Adjusted* ORs (95% CI) for probability of negative rapport building with medium and high versus low burnout: ▶ Medium: 1.85 (1.31, 2.61), p=0.001 ▶ High: 2.06 (1.58, 2.86), p<0.001 *Adjusted for patient health insurance, visit length, physician gender, physician international medical school graduate (IMG) status, interaction between IMG status and burnout

Continued

**Table 1** Continued

		Quality of care outcomes					
Author(s)	Study population	Description of sample	Burnout measure	Quality of care measure	Medical errors (ME)	Patient satisfaction (PS)/quality of care (QoC)	Communication/attitudes
Shanafelt <i>et al.</i> <sup>38</sup> USA	American surgeons who were members of the American College of Surgeons who permitted email correspondence. Response rate: 32%	n=7905 physicians Specialties: General: 41% Cardiothoracic: 6% Colorectal: 4% Otolaryngology: 5% Obstetrics/gynaecology: 1% Oncological: 5% Paediatric: 2% Plastic: 4% Transplant: 2% Trauma: 4% Urological: 4% Vascular: 6% Other: 6% Male: 87% Female: 13% Median age (IQR): 51 years (43, 59)	22-item MBI-Human Services Survey	Physician report: Response to: 'Are you concerned you have made any major medical error in the last 3 months?' Psychometric properties not tested.	ORs (95% CI) for perceived medical error with MBI dimensions: ▲ EE=1.048 (1.042, 1.055), p<0.0001 ▲ DP=1.109 (1.096, 1.122), p<0.0001 ▲ PA=0.965 (0.955, 0.975), p<0.0001		
Shirom <i>et al.</i> <sup>39</sup> Israel	Physicians from four health plans specialising in either: ophthalmology, dermatology, otolaryngology, gynaecology (community based), general surgery, cardiology (hospital based). 50% random probability sample drawn from each speciality. Response rate: 63%	n=890 physicians Male: 80% Female: 20% Median age: 52 years	12-items from the Shirom-Melamed Burnout Measure with three subscales: (1) physical fatigue, (2) cognitive weariness, and (3) emotional exhaustion.	Physician report: Physicians completed a 15-item version of the modified SERVQUAL. 5-point Likert scale from 1= 'to a very small extent' to 5= 'to a very large extent'. Psychometric properties of the modified version not tested.	Structural equation model examining relationships of autonomy, burnout and QoC: ▲ Relationship between global burnout and QoC not significant (β=-0.12, p>0.05) ▲ EE exhaustion negatively related to QoC (β=-.40, p<0.05)		
Travado <i>et al.</i> <sup>34</sup> Italy, Spain, Portugal	Physicians recruited from cancer centres of three hospitals: two general hospitals with a cancer ward and one cancer hospital. Convenience sample Response rate: Not reported.	n=125 physicians Years of practice: 15±9.4 years Male: 47% Female: 54% Mean age: 42±9.7 years	22-item MBI-Human Services Survey Used Maslach and Jackson <sup>6</sup> cut-off scores for no/low burnout, intermediate and high burnout.	Physician report: Communication skills assessed using two scales: (1) Self-Confidence in Communications Skills (SCSS). 12-item scale rating ability to communicate and manage a series of clinical situations. (2) Expected Outcomes of Communication (EOC). 23-item scale assessing extent to which physician perceives result of communication is positive or negative. Psychometric testing not completed. <sup>56</sup>	Correlations between MBI burnout dimensions and SCSS ▲ EE: r=-0.03, not significant ▲ DP: r=-0.08, not significant ▲ PA: r=0.37, p<0.01 Negative EOC ▲ EE: r=-0.21, p<0.05 ▲ DP: r=-0.25, p<0.01 ▲ PA: r=0.28, p<0.01 Positive EOC ▲ EE: r=0.01, not significant ▲ DP: r=0.34, p<0.01 ▲ PA: r=-0.28, p<0.01		

Continued

Table 1 Continued

Author(s)	Quality of care outcomes			
	Study population	Description of sample	Burnout measure	Quality of care measure
Weigl et al. <sup>66</sup> Germany	Physicians working in one academic children's hospital who were providing patient care. Response rate: 74%	n=88 physicians Years of practice: 8±6.7 years Male: 47% Female: 53% Mean age: 37±8.6 years	Two subscales of the German version of the MBI-D: EE and DP. High burnout defined as Mean EE score >3.5 and Mean DP >2.5	Physician report: 2-item perceived quality of care measure: 'My workload frequently leads to reduced quality of work' and 'Adverse work conditions frequently lead to a loss of quality.' 5-point Likert scale from 1='not at all' to 5='a very great extent'. Psychometric properties not tested for the two items taken from the German version of the MBI.
Wen et al. <sup>67</sup> China	Physicians practicing in one of 46 hospitals in 10 provinces n=12 tertiary hospitals n=9 secondary hospitals n=25 primary hospitals In the secondary and tertiary hospitals, physicians were selected from >10 clinical departments with >10 people in the age groups: <30 years, 30-39 years, 40-49 years and >50 years. Response rate: 89%	n=1607 total physicians n=192 physicians from primary hospitals n=354 physicians from secondary hospitals n=991 physicians from tertiary hospitals <i>Primary hospital</i> Male: 54% Female: 46% Mean age: 37±9.9 years Education: ≤high school: 17% Some college: 47% Bachelors' degree: 35% ≥Master's degree: 1.0% <i>Secondary hospital</i> Male: 53% Female: 47% Mean age: 36±9.4 years Education: ≤high school: 4% Some college: 17% Bachelors' degree: 73% ≥Master's degree: 6% <i>Tertiary hospital</i> Male: 61% Female: 39% Mean age: 36±8.0 years Education: ≤high school: 1% Some college: 3% Bachelors' degree: 46% ≥Master's degree: 51%	Used 15-item Chinese version of the MBI-General Survey Respondents grouped into three categories: (1) no burnout symptoms; (2) some burnout symptoms; (3) serious burnout symptoms.	Adjusted* ORs (95% CI) for probability of low QoC with MBI dimensions (low vs high): ▶ EE=0.75 (0.08, 1.42), p<0.05 ▶ DP=0.17 (-0.45, 0.80), not significant *Adjusted for gender, professional tenure, clinical work environment, career stage/position
			Physicians were asked if they had made any of the following medical errors: (1) patient was harmed, (2) medication error, (3) treatment delayed and (4) incomplete or incorrect item in the patient record. Psychometric properties not tested.	Adjusted* ORs (95% CI) for probability of any medical error with no burnout symptoms group as reference: ▶ Some burnout symptoms: 1.46 (1.13, 1.89) ▶ Serious burnout symptoms: 2.28 (1.63, 3.17) *Adjusted for sex, workload and hospital type

Continued

Table 1 Continued

Author(s)	Study population	Description of sample	Burnout measure	Quality of care measure	Medical errors (ME)	Quality of care outcomes	
						Patient satisfaction (PS)/quality of care (QoC)	Communication/attitudes
Weng <i>et al</i> <sup>55</sup> Taiwan	Physicians working in two hospitals. Patients of participating physicians. Physician response rate: Not reported. Patient response rate: 78%	n=110 internists Male: 85% Female: 15% Mean age: 41±6.9 years n=2872 patients Male: 59% Female: 41%	MBI-Human Services Survey	Patient report: Patient satisfaction assessed with two questions: 'I am satisfied with the care provided by my doctor' and 'I would recommend this doctor to my friends and family members'. Single item from the CSQ <sup>52</sup> General Satisfaction subscale not validated for individual use. Single item about recommendation was correlated with EUROPEP patient satisfaction questionnaire. <sup>53</sup>		Correlation between MBI burnout dimensions and PS: ▲ EE: not significant ▲ DP: negative relationship (p<0.01) ▲ PA: not significant	

DP, depersonalisation; EE, emotional exhaustion; PA, personal accomplishment.

interactions, Ratanawongsa *et al*<sup>33</sup> assessed the interactions by employing the Roter Interaction Analysis System (RIAS).<sup>54</sup> RIAS is a validated method of categorising these interactions into three categories related either to content, affection or process.<sup>55</sup> There is evidence that there is an association between the content and the socio-emotional nature of the interactions as categorised using the RIAS and patient satisfaction.<sup>54 55</sup>

Travado *et al*<sup>34</sup> examined the association between burnout and communication using two measures: the Self-Confidence in Communications Skills and the Expected Outcomes of Communication.<sup>56</sup> In their article, Parle and colleagues<sup>56</sup> note that exploration of the psychometric properties of both measures were being conducted but were not yet completed. Both were developed to understand the communication skills of physicians working with cancer patients.

#### Study outcomes: burnout and quality of care

In this subsection, we report about the quality of care outcomes from the included studies (table 1). This review of outcomes begins by describing the findings regarding the association between burnout and patient safety (ie, medical errors). It is followed by reporting of the acceptability outcomes as measured by patient satisfaction/perceived quality of care and physician communication/attitudes.

#### Outcomes: burnout and medical errors

Table 1 contains the outcomes reported by the included papers. In terms of findings for the association between burnout and medical errors, there was a consistently significant relationship between burnout and medical errors among four papers focusing on this relationship.<sup>5 30 37 38</sup> Shanafelt *et al*<sup>38</sup> reported significantly higher odds of a major medical error during the past 3 months among physicians with higher EE and DP but lower odds among physicians with higher PA. Hayashino *et al*<sup>30</sup> also observed significant associations between a major medical error during the past 12 months and higher levels of EE and DP; however, the relationship with PA was not significant. Klein *et al*<sup>5</sup> reported significant associations between high burnout and diagnostic error, therapeutic error, suboptimal psychosocial care, suboptimal diagnosis and treatment and suboptimal quality assurance. Wen *et al*<sup>37</sup> found higher odds of medical errors among physicians with either some or serious burnout symptoms as opposed to no burnout symptoms.

The one paper<sup>29</sup> that assessed errors based on chart audits did not find a significant relationship between burnout and medical errors. However, it should be noted that this study focused on treatment for a subgroup of patients with chronic disorders that included diabetes and/or hypertension.

#### Outcomes: burnout and patient satisfaction/quality of care

Among the four studies that examined the relationship between burnout and patient satisfaction/quality of care,

three studies observed a significant relationship between patient satisfaction/quality of care and either burnout or at least one dimension of burnout.<sup>31–33 35</sup> The one study<sup>33</sup> that combined the MBI EE and PA dimensions to create a single burnout score did not find a significant relationship between the score and patient satisfaction. Because it used only two subscales and one of them was PA rather than DP, it is not clear regarding the extent to which their choice of subscales was consistent with the other measures of burnout.

Among the three studies that reported separate MBI dimensions, there seemed to be a consistent observation that high DP is significantly related to lower patient satisfaction.<sup>31 32 35</sup> However, the significance of the association between EE and patient satisfaction varied among studies; Anagnostopoulos *et al*<sup>31</sup> reported a significant correlation, but Weng *et al*<sup>35</sup> did not.

At the same time, Shirom *et al*<sup>39</sup> described a significantly negative relationship between high EE and physician perceived quality of care. Weigl and colleagues<sup>36</sup> also found a significant negative relationship with EE but did not find a significant relationship between DP and physician perceived quality of care.

#### *Outcomes: burnout and communication/attitudes*

Travado *et al*<sup>34</sup> found a significantly positive relationship between PA and self-confidence in communication skills as well as with negative expected outcomes of communication. They also observed a significantly negative association between PA and positive expected outcomes of communication. In addition, Ratanawongsa *et al*<sup>33</sup> reported a higher probability of negative rapport with medium and high burnout.

## DISCUSSION

This systematic literature review identified 12 studies of which 10 had a moderate risk of bias and two had a high risk of bias. The results of these physician burnout studies show that patient safety has been primarily measured by examining medical errors. The acceptability outcomes have been captured using two groups of indicators that measure patient satisfaction/perceived quality of care and physician communication/attitudes towards patients. The majority of these studies examined the relationship between burnout and acceptability. Among the acceptability-related quality of care outcomes, the focus has been on patient satisfaction/perceived quality of care.

The results of four of the five included studies that reported on the relationship between burnout and medical errors suggest there is evidence that burnout is associated with physician self-perceived medical errors and suboptimal care. However, there is equivocal evidence that specific dimensions of burnout are related to the acceptability dimension of quality of care as measured by patient satisfaction, perceived quality of care or physician communication/attitudes. Thus, the current body of evidence suggests there is moderate

evidence for the association between burnout and safety aspects of healthcare, whereas the evidence is weaker for the patient-related acceptability aspects of quality.

### Strengths and limitations of interpreting the literature

One of the important questions raised by burnout studies in general is highlighted by Klein *et al*'s<sup>5</sup> and Shirom *et al*'s<sup>39</sup> use of non-MBI scales. Klein and colleagues<sup>5</sup> used the CBI, while Shirom *et al*<sup>39</sup> used the SMBM. One of the criticisms that the separate developers of these two scales raise is that the MBI does not fully assess burnout.<sup>39 40</sup> Rather, both groups argue that fatigue and exhaustion are fundamental to the definition of burnout.<sup>39 40</sup> However, this emphasis on exhaustion may be reflected in the fact that EE is the most widely studied of the MBI dimensions.<sup>57</sup> This would argue for the assessment of this dimension in studies of burnout and the individual reporting of it.

Another limitation of these studies was the reliance on physician self-report data for the assessment of medical errors. The self-report could be influenced by a number of factors including recall bias and social desirability. There is a potential additional bias introduced if self-report is used for both the outcome and the problem.<sup>58</sup> The presence of burnout could also influence perceptions. For example, Fahrenkopf *et al*<sup>59</sup> observed a discrepancy between the results of chart audits and physician self-report; those with higher burnout scores reported higher numbers of medical errors than the chart audits would suggest.

An alternative to self-report would be observational data. However, watching physicians while they practice could lead to a Hawthorne effect. Another alternative would be to review medical records to identify errors. However, this relies on the accuracy of the records. Also, it is not clear what types of medical errors should be assessed—major errors leading to an adverse event or any medical error regardless of outcome? In their study, Fahrenkopf *et al*<sup>59</sup> used a standardised method to abstract information from charts and trained reviewers to categorise the errors into groups: (1) preventable adverse event, (2) non-preventable adverse event, (3) potential adverse event and (4) error with little potential for harm. Further work could examine how physicians define errors as well as the reliability of error self-report. In addition, to improve the comparability of outcomes, future studies could incorporate and report severity of medical error scores.

There was a diverse set of measures used in the studies that focused on patient satisfaction and quality of care. They varied in which outcomes were measured and how they were measured. In addition, the majority of the studies did not use validated outcome measures. For example, perceived quality of care was assessed using a variety of measures that ranged from two items for which the psychometric properties were not tested to a scale designed to assess service quality on six dimensions. Thus, it is difficult to discern the extent to which the study results could be attributed to the differences in the dimensions assessed. Further exploration along this line of inquiry could be

undertaken to understand the aspects of satisfaction and perceived quality of care that are significantly associated with burnout.

An additional limitation of the existing body of literature is the reliance on cross-sectional study designs. Cross-sectional design limits conclusions regarding causality. Cross-sectional data do not distinguish the sequence of conditions. For example, did burnout cause decreased quality of care? Or, did decreased quality of care cause burnout? At best, the cross-sectional data used in these studies can only be used to determine that there is a relationship. At the same time, there is evidence from studies that have used longitudinal data to examine burnout and medical errors among residents that there is a causal relationship such that burnout causes errors.<sup>60</sup> However, the longitudinal data that contribute to the strength of West *et al's* study<sup>60</sup> are potentially weakened by the self-reported medical errors.

Finally, only two studies<sup>5,29</sup> described the population from which the study sample was drawn. Thus, it is difficult to determine whether there was a difference between the study participants and non-participants. To aid in the interpretation of the results (ie, the generalisability), it would be useful for future studies to report this type of information.

### Strengths and limitations of the search strategy

Although six databases were used in the search, articles that did not appear in any of the databases would have been missed. To decrease the possibility of this occurring, we employed a broad scope in development of the search terms for each database and followed this with a hand search of included articles. Another potential limitation is the fact that the search focused on articles published in English-language journals. However, despite the English-language constraint, the identified studies originated in European, Middle Eastern, North American and Asian countries. This indicates that although the research was not conducted in countries where English is the first language, at least some of these researchers publish in English-language journals. Finally, there is also a potential limitation associated with focusing on published peer-reviewed articles. In doing so, we may be subject to publication bias.<sup>61</sup> At the same time, the quality of the grey literature has been questioned, because it is not necessarily subject to critical assessment prior to being published.<sup>62</sup> As a result, unpublished studies may be of lower quality and have greater risk of bias in their study designs.

### CONCLUSIONS

The focus on quality related to direct care can highlight additional ways that physician burnout affects the health-care system. These results contribute evidence about whether the effects of physician burnout are limited to physicians or whether consequences of physician burnout are more extensive. They also can help to inform decisions about how to improve patient care by addressing physician burnout. That is, decisions can be informed

when confronting a question of how to improve quality of patient care. There are a number of ways in which this may be done through investment in capital such as new technologies. The results of this systematic review suggest that an alternative investment could be in human resources as represented by physician staff.

The results of this systematic literature review suggest that there is moderate evidence that burnout is associated with safety-related quality of care. Because of the variability in the way patient acceptability-related quality of care was measured and the inconsistency in study findings, the evidence supporting the relationship between burnout and patient acceptability-related quality of care is less strong. Future research evaluating burnout interventions for physicians could consider looking at safety-related quality of care to assess the effectiveness of these interventions. Continued work looking at the relationship between dimensions of acceptability-related quality of measures and burnout is warranted.

**Correction notice** This paper has been amended since it was published Online First. Owing to a scripting error, some of the publisher names in the references were replaced with 'BMJ Publishing Group'. This only affected the full text version, not the PDF. We have since corrected these errors and the correct publishers have been inserted into the references.

**Contributors** CSD led the conception, design, data acquisition, analysis and interpretation of the data; she also led the writing of the overall manuscript. DL collaborated on the design, data acquisition and analysis; he contributed to the writing of the overall manuscript and led the writing of the Methods section. SB collaborated on the design and data acquisition and contributed to the writing of the manuscript. LT collaborated on the data acquisition and analysis. All authors read and approved the final manuscript. All authors are guarantors of the final manuscript.

**Competing interests** None declared.

**Patient consent** This study did not involve human subjects.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** There is no additional unpublished data from this study.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

### REFERENCES

- Allegra CJ, Hall R, Yothers G. Prevalence of burnout in the u.s. Oncology community: results of a 2003 survey. *J Oncol Pract* 2005;1:140–7.
- Arigoni F, Bovier PA, Sappino AP. Trend of burnout among Swiss doctors. *Swiss Med Wkly* 2010;140:w13070.
- Elit L, Trim K, Mand-Bains IH, *et al*. Job satisfaction, stress, and burnout among Canadian gynecologic oncologists. *Gynecol Oncol* 2004;94:134–9.
- Embriaco N, Azoulay E, Barrau K, *et al*. High level of burnout in intensivists: prevalence and associated factors. *Am J Respir Crit Care Med* 2007;175:686–92.
- Klein J, Grosse Frie K, Blum K, *et al*. Burnout and perceived quality of care among German clinicians in surgery. *Int J Qual Health Care* 2010;22:525–30.
- Maslach C, Jackson SE. The measurement of experienced burnout. *J Organ Behav* 1981;2:99–113.

7. Maslach C, Schaufeli WB, Leiter MP. Job burnout. *Annu Rev Psychol* 2001;52:397–422.
8. Dewa CS, Jacobs P, Thanh NX, et al. An estimate of the cost of burnout on early retirement and reduction in clinical hours of practicing physicians in Canada. *BMC Health Serv Res* 2014;14:254.
9. Sharma A, Sharp DM, Walker LG, et al. Stress and burnout in colorectal and vascular surgical consultants working in the UK National Health Service. *Psychooncology* 2008;17:570–6.
10. Siu C, Yuen SK, Cheung A. Burnout among public doctors in Hong Kong: cross-sectional survey. *Hong Kong Med J* 2012;18:186–92.
11. Asai M, Morita T, Akechi T, et al. Burnout and psychiatric morbidity among physicians engaged in end-of-life care for cancer patients: a cross-sectional nationwide survey in Japan. *Psychooncology* 2007;16:421–8.
12. Dewa CS, Loong D, Bonato S, et al. How does burnout affect physician productivity? A systematic literature review. *BMC Health Serv Res* 2014;14:325.
13. World Health Organization. *Quality of care: a process for making strategic choices in health systems*. Geneva, 2006.
14. Institute of Medicine. *Crossing the Quality Chasm: a New Health System for the 21st Century*. Washington, D.C, 2001.
15. Roberts JS. Chapter 1 Quality Health Care: Its Definitions and Evaluation. Huges EFX, editor. *Perspectives on Quality in American Health Care*. Washington, DC: mcgraw-hill, Inc, 1988.
16. Wallace JE, Lemaire JB, Ghali WA. Physician wellness: a missing quality Indicator. *Lancet* 2009;374:1714–21.
17. de Jong MA, Nieuwenhuijsen K, Sluiter JK. Common mental disorders related to incidents and behaviour in physicians. *Occup Med* 2016;66:506–13.
18. Hall LH, Johnson J, Watt I, et al. Burnout, and patient safety: a systematic review. *PLoS One* 2016;11:e0159015.
19. Williams ES, Skinner AC. Outcomes of physician job satisfaction: a narrative review, implications, and directions for future research. *Health Care Manage Rev* 2003;28:119–39.
20. Salyers MP, Bonfils KA, Luther L, et al. The relationship between professional burnout and quality and safety in healthcare: a meta-analysis. *J Gen Intern Med* 2017;32:475–82.
21. Angerer P, Weigl M. Physicians' psychosocial work conditions and quality of care: a literature review. *Professions and Professionalism* 2015;5:1–14.
22. Lee RT, Seo B, Hladkyj S, et al. Correlates of physician burnout across regions and specialties: a meta-analysis. *Hum Resour Health* 2013;11:48.
23. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.
24. McGowan J, Sampson M, Salzwedel DM, et al. PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. *J Clin Epidemiol* 2016;75:40–6.
25. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37–46.
26. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol* 2007;36:666–76.
27. von Elm E, Altman DG, Egger M, et al. . The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med* 2007;4:e296.
28. Lagerveld SE, Bültmann U, Franche RL, et al. Factors associated with work participation and work functioning in depressed workers: a systematic review. *J Occup Rehabil* 2010;20:275–92.
29. Rabatin J, Williams E, Baier Manwell L, et al. Predictors and outcomes of burnout in Primary Care Physicians. *J Prim Care Community Health* 2016;7:41–3.
30. Hayashino Y, Utsugi-Ozaki M, Feldman MD, et al. Hope modified the association between distress and incidence of self-perceived medical errors among practicing physicians: prospective cohort study. *PLoS One* 2012;7:e35585.
31. Anagnostopoulos F, Liolios E, Persefonis G, et al. Physician burnout and patient satisfaction with consultation in primary health care settings: evidence of relationships from a one-with-many design. *J Clin Psychol Med Settings* 2012;19:401–10.
32. Halbesleben JR, Rathert C. Linking physician burnout and patient outcomes: exploring the dyadic relationship between physicians and patients. *Health Care Manage Rev* 2008;33:29–39.
33. Ratanawongsa N, Roter D, Beach MC, et al. Physician burnout and patient-physician communication during primary care encounters. *J Gen Intern Med* 2008;23:1581–8.
34. Travado L, Grassi L, Gil F, et al. Physician-patient communication among Southern European cancer physicians: the influence of psychosocial orientation and burnout. *Psychooncology* 2005;14:661–70.
35. Weng HC, Hung CM, Liu YT, et al. Associations between emotional intelligence and doctor burnout, job satisfaction and patient satisfaction. *Med Educ* 2011;45:835–42.
36. Weigl M, Schneider A, Hoffmann F, et al. Work stress, burnout, and perceived quality of care: a cross-sectional study among hospital pediatricians. *Eur J Pediatr* 2015;174:1237–46.
37. Wen J, Cheng Y, Hu X, et al. Workload, burnout, and medical mistakes among physicians in China: a cross-sectional study. *Biosci Trends* 2016;10:27–33.
38. Shanafelt TD, Balch CM, Bechamps G, et al. Burnout and medical errors among American surgeons. *Ann Surg* 2010;251:995–1000.
39. Shirom A, Nirel N, Vinokur AD. Overload, autonomy, and burnout as predictors of physicians' quality of care. *J Occup Health Psychol* 2006;11:328–42.
40. Kristensen TS, Borritz M, Villadsen E, et al. The copenhagen burnout inventory: a new tool for the assessment of burnout. *Work Stress* 2005;19:192–207.
41. Lundgren-Nilsson Åsa, Jonsdottir IH, Pallant J, et al. Internal construct validity of the Shirom-Melamed Burnout Questionnaire (SMBQ). *BMC Public Health* 2012;12:12:1.
42. Dolan ED, Mohr D, Lempa M, et al. Using a single item to measure burnout in primary care staff: a psychometric evaluation. *J Gen Intern Med* 2015;30:582–7.
43. Shirom A, Melamed S. A comparison of the construct validity of two burnout measures in two groups of professionals. *Int J Stress Manag* 2006;13:176–200.
44. Qiao H, Schaufeli WB. The convergent validity of four Burnout measures in a chinese sample: a confirmatory Factor-Analytic approach. *Appl Psychol* 2011;60:87–111.
45. Winwood PC, Winefield AH. Comparing two measures of Burnout among Dentists in Australia. *Int J Stress Manag* 2004;11:282–9.
46. Kaldjian LC, Jones EW, Wu BJ, et al. Reporting medical errors to improve patient safety: a survey of physicians in teaching hospitals. *Arch Intern Med* 2008;168:40–6.
47. Davis DA, Mazmanian PE, Fordis M, et al. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA* 2006;296:1094–102.
48. Hall W, Violato C, Lewkonja R, et al. Assessment of physician performance in Alberta: the physician achievement review. *CMAJ* 1999;161:52–7.
49. Parasuraman A, Zeithaml VA, Berry LL. SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing* 1988;64:12–40.
50. Scardina SA. SERVQUAL: a tool for evaluating patient satisfaction with nursing care. *J Nurs Care Qual* 1994;8:38–46.
51. Asubonteng P, McCleary KJ, Swan JE. SERVQUAL revisited: a critical review of service quality. *Journal of Services Marketing* 1996;10:62–81.
52. Baker R. Development of a questionnaire to assess patients' satisfaction with consultations in general practice. *Br J Gen Pract* 1990;40:487–90.
53. Kersnik J. Patients' recommendation of doctor as an Indicator of patient satisfaction. *Hong Kong Med J* 2003;9:247–50.
54. Roter D. An exploration of health education's responsibility for a partnership model of client-provider relations. *Patient Educ Couns* 1987;9:25–31.
55. Roter DL, Stewart M, Putnam SM, et al. Communication patterns of primary care physicians. *JAMA* 1997;277:350–6.
56. Parle M, Maguire P, Heaven C. The development of a training model to improve health professionals' skills, self-efficacy and outcome expectancies when communicating with cancer patients. *Soc Sci Med* 1997;44:231–40.
57. Schaufeli WB, Leiter MP, Maslach C. Burnout: 35 years of research and practice. *Career Development International* 2009;14:204–20.
58. Podsakoff PM, MacKenzie SB, Lee JY, et al. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J Appl Psychol* 2003;88:879–903.
59. Fahrenkopf AM, Sectish TC, Barger LK, et al. Rates of medication errors among depressed and burnt out residents: prospective cohort study. *BMJ* 2008;336:488–91.
60. West CP, Tan AD, Habermann TM, et al. Association of resident fatigue and distress with perceived medical errors. *JAMA* 2009;302:1294–300.
61. Nissen SB, Magidson T, Gross K, et al. Publication bias and the canonization of false facts. *Elife* 2016;5.
62. Martin JL, Pérez V, Sacristán M, et al. Is grey literature essential for a better control of publication bias in psychiatry? An example from three meta-analyses of schizophrenia. *Eur Psychiatry* 2005;20:550–3.





63. Poulton BC. Use of the consultation satisfaction questionnaire to examine patients' satisfaction with general practitioners and community nurses: reliability, replicability and discriminant validity. *Br J Gen Pract* 1996;46:26–31.
64. Roter D, Larson S. The Roter Interaction Analysis System (RIAS): utility and flexibility for analysis of medical interactions. *Patient Educ Couns* 2002;46:243–51.