European Working Time Directive and doctors' health. A Systematic review of the available epidemiological evidence.

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-004916
Article Type:	Research
Date Submitted by the Author:	24-Jan-2014
Complete List of Authors:	Rodriguez-Jareno, Maria-Cruz; Universitat de Girona, 2. Department of Medical Sciences Demou, Evangelia; University of Glasgow, Institute of Health and Wellbeing Vargas, Sergio; Universitat Pompeu Fabra, Centre for Research in Occupational Health Sanati, Kaveh; University of Glasgow, Institute of Health and Wellbeing Skerjanc, Alenka; University Medical Centre, 6. Clinical Institute of Occupational, Traffic and Sports Medicine Reis, Pedro; 7. College of Occupational Medicine, Portuguese Medical Association Helimaki-Aro, Ritva; 8. Helsinki City Occupational Health Centre, Macdonald, Ewan; University of Glasgow, Institute of Health and Wellbeing Serra, Consol; Universitat Pompeu Fabra, Centre for Research in Occupational Health
Primary Subject Heading:	Health policy
Secondary Subject Heading:	Medical management, Research methods, Health services research, Health policy, Evidence based practice
Keywords:	EPIDEMIOLOGY, GENERAL MEDICINE (see Internal Medicine), HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

European Working Time Directive and doctors' health. A Systematic review of the available epidemiological evidence.

Running title: Long working hours and doctors' health

***Corresponding author**

Dr. Evangelia Demou
Healthy Working Lives Group
Institute of Health and Wellbeing
University of Glasgow
Glasgow, G12 8RZ
Tel: +44 141 330 3559
E-mail: evangelia.demou@glasgow.ac.uk

Maria Cruz Rodriguez-Jareño[1,2], Evangelia Demou[3,*], Sergio Vargas-Prada[4], Kaveh A Sanati[3,5], Alenka Škerjanc[6], Pedro G. Reis[7], Ritva Helimäki-Aro[8], Ewan B. Macdonald[3], Consol Serra[1,4,9,10], on behalf of the UEMS Section of Occupational Medicine[11].

1. Catalan and Spanish Societies of Occupational Medicine, Barcelona, Spain.
2. Department of Medical Sciences. School of Medicine. Universitat de Girona. Girona. Spain
3. Institute of Health and Wellbeing, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom
4. CiSAL - Centre for Research in Occupational Health, Universitat Pompeu Fabra, Barcelona, Spain.
5. Occupational Health Department, Epsom & St Helier University Hospitals NHS Trust, Surrey, United Kingdom
6. Clinical Institute of Occupational, Traffic and Sports Medicine, University Medical Centre, Ljubljana, Slovenia.
7. College of Occupational Medicine, Portuguese Medical Association, Lisbon, Portugal.
8. Helsinki City Occupational Health Centre, Helsinki, Finland.
9. Occupational Health Service, Parc de Salut MAR, Barcelona, Spain.
10. CIBER of Epidemiology and Public Health. Spain.
11. UEMS Section of Occupational Medicine. http://www.uems-occupationalmedicine.org

Word Count: 4,667

ABSTRACT

Objective: To summarise the available scientific evidence on the health effects of exposure to working beyond the limit number of hours established by the European Working Time Directive (EWTD) on physicians.

Design: A systematic literature search was conducted in PubMed and Embase. Study selection, quality appraisal and data extraction were carried out by independent pairs of researchers using pre-established criteria.

Setting: Physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.

Participants: The total number of participants was 14,338.

Primary and secondary outcome measures: Health effects classified under the International Classification of Diseases (ICD-10).

Results: Over 3,000 citations and 110 full articles were reviewed. From these, eleven studies of high or intermediate quality carried out in North America, Europe and Japan met the inclusion criteria. Six studies included medical residents, junior doctors or house officers and the five others included medical specialists or consultants, medical, dental and general practitioners and hospital physicians. Evidence of an association was found between percutaneous injuries and road traffic accidents with extended hours/day or very long working weeks (LWH). The evidence was insufficient for mood disorders and general health. No studies on other health outcomes were identified.

Conclusions: LWH could increase the risk of percutaneous injuries and road traffic accidents, and possibly other incidents at work through the same pathway. While associations are clear, the existing evidence does not allow for an established causal or "dose-response" relationship between LWH and incidents at work, or for a threshold number of extended hours above which there is a significantly higher risk and the hours physicians could work and remain safe and healthy. Policy makers should consider safety issues when working on relaxing EWTD for doctors.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- The first systematic review, to our knowledge, on the effects of Long Working Hours (LWH) on physicians' health and safety.
- A systematic literature search conducted in PubMed and Embase with over 3,000 citations and 110 full articles reviewed.
- Eleven studies of high or intermediate quality carried out in North America, Europe and Japan, involving physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.
- The findings of this review confirmed that long working hours are associated with both an increased risk of needlestick injuries and road traffic accidents. Evidence was assessed as low or insufficient for mood disorders and general health.
- This review uncovered the lack of literature on the effect of long working hours on the general health of doctors, a topic that has important health ramifications not only for physicians but for patients too.

KEY WORDS: Working time, "Physicians"[MeSH], "Morbidity"[MeSH], "Occupational Injuries"[MeSH], Systematic review.

INTRODUCTION

The European Working Time Directive (EWTD),[1] applicable to all occupations across the EU, requires a maximum working week of 48 hours and establishes rest periods.[2, 3] Since it came into force in healthcare in 2009, it has been associated with concerns about the provision of health services including continuity of care, lower staffing levels, introduction of shift working, a reduction in training time and the adequate supervision of junior doctors.[4, 5] The large inter-country variations in physicians working hours as well as the mandatory controls over work hours across occupations raises further questions as to the basis of selecting and setting these limits and restrictions.[6] There is evidence of variable compliance to the EWTD within health care across Europe.[4, 5] The medical profession is also increasingly feminised with potential implications for women of childbearing age, which may coincide with the time in the profession where long working hours (LWH) are more likely to happen (i.e. during residency). Constraint on public-sector finances and increasing health care demands, have stimulated the debate on physicians' working hours,[7] with some evidence showing that the quality of health care is positively correlated with the number of physicians available to deliver that care.[8]

The Union of European Medical Specialists (UEMS) was created in 1958 to represent medical specialists in the EU, promote a high standard of training and practice at European level and facilitate the free movement of physicians across European countries.[9] The Occupational Medicine section was created in 1997 and is involved in issues related to specialist training, professional practice and research;[10] elaborating and publishing reference documents,[11] conducting and contributing to surveys,[12-14] and working in partnership with other key European groups.

The trigger for this project was a request by the Council of the UEMS to its Occupational Medicine Section for a position statement on the EWTD and possible health consequences to physicians of a potential removal of this professional group from the current EWTD 48 hours per week limit. An evidence based approach was taken by the Section and a systematic review of the literature was undertaken.

LWH have been associated with several adverse effects on workers' health, such as hypertension, cardiovascular disease, stress, depression, musculoskeletal disorders, chronic infections, diabetes, general health complaints and all-cause mortality in a variety of occupational groups.[15-23] In physicians, there is some evidence suggesting

that LWH can result in impacts such as stress, depression, burnout, injuries, fatigue and sleep deprivation,[24-26] and overwork has been considered one of the most stressful features of physicians' work.[24] Some evidence exists of the beneficial effect of reducing the number of working hours on the quality of life of medical residents [27] and on burnout.[24]

However, the relationship between work and health is complex. There is a substantial body of evidence showing that worklessness is associated with poorer health, whereas work is generally good for health and well-being,[28, 29] provided that the work environment is reasonably acceptable and supportive.[30]

Occupational and non-occupational exposures may play an important role such as the pattern and distribution of working hours, breaks, and recovery periods. Also, psychosocial and organisational factors, such as workload, job control, managers and peer support, training opportunities, individual characteristics and attitudes may be important.[6, 31]

Finally, physicians are a highly qualified, devoted and motivated professional group. Despite other existing reviews of a more general occupational approach, a review focused on physicians is warranted.

The aim of this project was to systematically review whether long-working hours (LWH) -i.e. defined as more than the 48 hour per week limit imposed by the EWTD- are associated with health effects –i.e. classified under the International Classification of Diseases (ICD-10) [32]- in physicians, and to examine what these associations are. This review does not include the potential impact of LWH on patient care or physician training.

METHODS

The systematic review sought to respond to the following research question: are LWH associated with health effects in physicians? For the purpose of this review, exposure to LWH was defined as working for more than 48 hours per week.[1] Health outcomes included any disease as defined in the ICD-10 [32] and work related injuries. If self-assessed tools on health status were used, only those validated were considered. Burnout, stress and outcomes such as individual symptoms, signs or biological markers (blood pressure, ECG, etc.) were not considered. Physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.

Study identification

Electronic searches were carried out using PubMed and EMBASE as search engines (December 2011). Our search strategy was similar in both databases and consisted of a combination of the following keywords and Mesh terms: night shift, morning shift, evening shift, afternoon shift, shift work, rotating shift, shift combination, shift duration or length, shift system, clockwise rotation, shift roster, extended shifts, night work, evening work, work schedule, work hours, starting time, early start, irregular working hours, direction of rotation, overwork, extended hours, shift rota, workload, work schedule tolerance, sleep deprivation, sleep disorders, chronobiology disorders, circadian rhythm, psychomotor performance, circadian disruption, vigilance, alertness, wakefulness, drowsiness, fatigue, insomnia, hypersomnolence, dyssomnia, eveningness, morningness, neurocognitive performance, concentration difficulties, arousal, health, morbidity, mortality, disease, illness, stress, strain, distress, accident, injur*, death, suicid*, education, medical, physician, medical staff, hospital, doctor, surgeon, house officers, medical school, surgery, surgical.

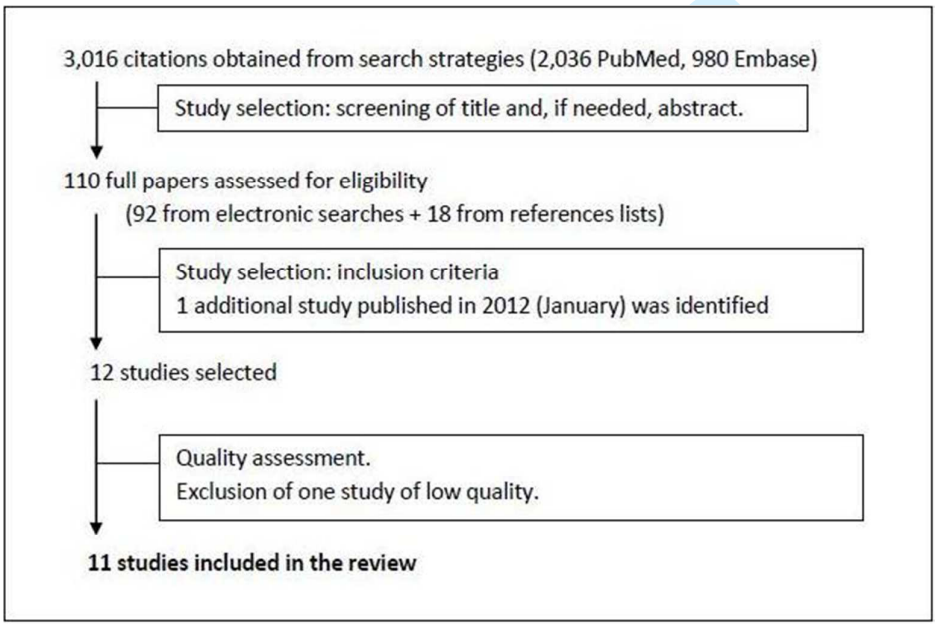
In addition, the references lists of articles selected for inclusion were carefully reviewed to identify additional studies.

Study selection

Studies were included if they measured the association between the exposure to LWH and health effects in physicians, used an observational epidemiological (i.e. cohort, case-control, cross-sectional) or experimental design, and were published in English, French, German, Italian, Slovenian or Spanish. Articles relating to on-call, night or shift work, but with unknown exposure or exposed to less than 48 working hours per week were excluded. They were also excluded if the working groups were other than physicians or addressed only other exposures or outcomes (i.e. patient safety, fatigue, sleep deprivation, social or family disruption).

A total of 2,036 citations were obtained from PubMed and 980 non-duplicated citations using Embase, yielding a total of 3,016 citations that were all screened by independent pairs of reviewers. All titles and, where necessary, abstracts were reviewed. Ninety two potentially suitable publications were identified from the electronic searches, and 18 further studies were identified from the references lists, yielding a total of 110 studies for all of which the full text was obtained and reviewed by independent pairs of reviewers. Disagreements within pairs were resolved by discussion and, where necessary, by a third reviewer who made the final decision. All authors participated as reviewers for screening the citations and full papers using well defined and pre-established criteria. Ninety nine articles were excluded at this stage. One further study was identified that was published after the search period. Finally, 12 papers were considered for quality assessment (Figure 1).

Figure 1. Results of the search strategy, using search engines on PubMed and EMBASE (December 2011) and screening of references lists of identified full papers, study selection and quality assessment.



The methodological quality of the 12 studies was assessed by independent pairs of reviewers. A standardized 16 point scale based on CONSORT and STROBE statements and adapted from a previous systematic review [33] was used. It includes 16 items grouped into 6 areas: 1) objectives, 2) study design, 3) target population and sample, 4) variables, 5) data sources, collection and measurement, and 6) statistical methods. Each item was rated as 1 (the requirement was met), 0.5 (the requirement was partially met) or 0 (the requirement was not or unclearly met). Disagreements within pairs of reviewers were resolved by consensus or, where necessary, by a third reviewer. For each of the 12 studies, a final score based on the sum of all items was assigned and the percentage was calculated based on a maximum score of 16. Study quality was rated as low, moderate or high if it scored less than 60%, between 60% and 79.9%, and 80% or more of the maximum score, respectively.

One article of low quality was excluded [34] and a total of 11 studies of intermediate or high methodological quality were included for the purpose of this review (figure 1).

Data extraction

Selected information was extracted from each paper, including publication year, country of origin, study design, setting, study population, sample size, response rate, measure of working hours, definition used for LWH, health outcomes and their measurement, and main results on point risk estimates or frequencies of health outcomes, and their corresponding 95% confidence interval (95%CI), and whether the analyses were adjusted for potential confounders.

Evidence synthesis

To summarise the results on the relationship between LWH and health effects, levels of evidence synthesis was performed. This was based on the methodological quality, study design and the consistency of the study outcomes. The following criteria were based on two previous relevant systematic reviews.[35, 36]

Strong evidence: consistent results in more than two studies of high quality.

Moderate evidence: consistent results in one high quality study and one intermediate, or between some studies of intermediate quality.

Insufficient evidence: identification of only one study or inconsistent results across studies;

Evidence of no association: consistent results of a non-association in two or more studies.

A measure of the possible magnitude of the association was attempted using the following criteria based on the association point estimate (RR = relative risk, rate ratio or odds ratio) when available or otherwise the comparison between two frequencies was used:

- high: RR > 3·00
- intermediate: RR = 1·50-3·00
- low: RR = 1·01-1·49
- no association: RR=1·00
- unclear

RESULTS

We identified 11 studies related to LWH and health effects in physicians that were eligible for inclusion. The outcome of the methodological quality assessment is given in table 1.

Methodological quality was appraised as high for 3 of the 11 studies, with scores ranging from 84 to 94%. The other 8 studies were considered as intermediate quality as their scores ranged from 63 to 78%. The majority of all included studies received positive scores on items describing the study objectives, design and the study population (items from areas 1, 2 and 3), as well as the description of exposure and outcome variables (item 4). However, although data sources and collection were well described, reliability and validity of exposure to LWH and health outcome measures had low scores (items 5b and 5c). The statistical methods were in general appropriate, however the confounding variables were not measured in several studies and, as can be seen, the studies did not describe how they addressed missing data (items 6a and 6b).

Study characteristics:

The characteristics of the included studies are described in table 2.

Table 1. Methodological quality appraisal of the included studies

Study ID			Ayas 2006 [37]	Barger 2005 [38]	Firth-Cozens 1987 [39]	Fisman 2007 [40]	Hayasaka 2007 [41]	Kirkcaldy 1997 [42]	Rosta 2011 [43]	Stamp 2005 [44]	Sundquist 2000 [45]	Varma 2012 [46]	Zahrai 2011 [47]
1	Objectives	Are the objectives or hypotheses of the research described in the paper?	1	1	1	1	1	1	1	1	1	1	1
2	Study design	Is the study design presented?	1	1	1	1	1	1	1	1	1	1	1
3a	Target population	Do the authors describe the target population they wanted to research?	1	1	1	1	1	1	1	1	1	1	1
3b	Sample	Was a random sample of the target population taken? AND was the response rate 60% or more?	0.5	0.5	1	0.5	0.5	0.5	0.5	1	1	1	0.5
3c	Sample	Is participant selection described?	1	1	1	1	0.5	1	0	1	1	1	1
3d	Sample	Is participant recruitment described, or referred to?	1	1	1	1	1	1	1	1	1	1	0
3e	Sample	Are the inclusion and/or exclusion criteria stated?	1	1	1	1	0.5	0.5	1	1	1	1	0.5
3f	Sample	Is the study sample described? (minimum description: sample size, gender, age and occupation)	1	1	0	1	1	1	0	0	0	1	1
3g	Sample	Is the number of participants at each stage of the study reported?	1	1	1	0.5	1	0.5	1	1	1	1	1
4	Variables	Are the measures of long working hours and the health outcome described?	1	1	0.5	1	1	1	1	1	1	1	1
5a	Data sources, collection	Do authors describe the source of their data (e.g., official registry, health survey) AND how the data were collected?	1	1	1	1	1	1	1	1	1	1	1
5b	Measurement	Was reliability of the measure(s) of long working hours mentioned or referred to?	0	0	0	1	0	0	0	1	0	1	0
5c	Measurement	Was the validity of the measure(s) of long working hours mentioned or referred to?	1	1	0	0	0	0	0	0	0	0	0
5d	Measurement	Were health outcomes assessed by objective measures or validated self-reporting instruments?	0	1	1	1	1	0	1	1	1	1	1

6a	Statistical methods	Were appropriate statistical methods used and described, including those for addressing confounders?	0.5	1	0.5	1	0.5	0.5	1	0.5	1	1	0.5
6b	Statistical methods	Were the numbers/% of participants with missing data for long working hours and the health outcome indicated AND If more than 20% of data in the primary analyses were missing, were methods used to address missing data?	0	0	0	0.5	0	0	0	0	0	1	0.5
Total score** (%)			12.0 (75)	13.5 (84)	11.0 (69)	13.5 (84)	11.0 (69)	10.0 (63)	10.5 (66)	12.5 (78)	12.0 (75)	15.0 (94)	11.0 (69)
Quality rate***			interm	high	interm	high	interm	interm	interm	interm	interm	high	interm

** Maximum score = 16; *** Quality rate (%): low=<60; intermediate (interm)=60-79; high=80-100.

Table 2. Characteristics of the included studies.

Study ID	Country	Design*	Setting	Participants and sample size (response rate)	Working hours	Health outcomes	Quality score (%)
Ayas 2006 [37]	United States	Cross-over	Hospital	Interns in postgraduate residency programs Sample size: 2,737 (80%)	Mean (SD) hrs 75.3:self-reported, monthly survey. Strong correlation (Pearson *r*=0.76; *p* 0.001) with hours worked (69.3) from work diaries of randomly selected subset of interns. Extended periods (>20 hrs/day) vs. non extended periods (first 12 hrs at work).	Self-reported percutaneous injuries	12.0 (75)
Barger 2005 [38]	United States	Cohort, prospective (1 year)	Hospital	Residents first postgraduate year (interns). Sample size: 2,554 (75%)	Average 70.7±26.0 hrs/week. Extended shift (≥ 24hrs) vs. non extended shift (<24 hours). Self-reported, validated.	Self-reported and documented motor vehicle crashes.	13,5 (84)
Firth Cozens 1987 [39]	United Kingdom	Cohort, prospective (1 year)	Hospital	Junior house officers Sample size: 170 (72%)	Mean number of hours per week = 90.6 hours (include on call). Self-reported.	GHQ-12 (case: score ≥2) and SCLDS, self-reported	11.0 (69)
Fisman 2007 [40]	United States	Cross-over	Hospital	Medical trainees vs. other HCW. Sample size: 109 vs. 241 (46%)	Median number of hours per week: medical trainees = 70; other HCW = 40 (p<0.001). Self-reported, high reliability.	Reported percutaneous injuries to employee health care service.	13.5 (84)
Hayasaka 2007 [41]	Japan	Cross-sectional	Hospital, clinics, other	Female physicians Sample size: 367 (63%)	Comparison of increasing number of hrs/week from ≤ 30 to >50, self-reported.	GHQ-30 (case: score ≥8), self-reported	11.0 (69)
Kirkcaldy 1997[42]	Germany	Cross-sectional	Not specified	Medical and dental practitioners. Sample size: 2,500 (not specified)	Long hours: mean 58.36 hrs/week Short hours: mean 38.17 hrs/week Self-reported.	Self-reported traffic accidents	10.0 (63)
Rosta 2011[43]	Germany	Cross-sectional	Hospital	Hospital physicians Sample size: 1,260 Germans and 562 Norwegians (58% and 65%,)	German vs. Norwegian physicians (%): Hours per day >9 = 58.8 vs. 26.7 60 hours on-call per month = 63.4 vs. 18.3 Self-reported.	Validated questionnaire on self-rated health	10.5 (66)
Stamp 2005 [44]	United States	Before-and-after	Hospital	Residents of general surgery Sample size: 28 (97%)	Changes of work patterns, after reduction from 90-110 to 78-80 work hrs/week.	SF-36, mental health; BDI, depression; self-reported	12.5 (78)
Sundquist 2000 [45]	Sweden	Cross-sectional	Primary care	General practitioners Sample size: 1,004 (72%)	Overtime defined as working at least 47 hrs/week. Self-reported.	Swedish SF-36, impaired mental health, self-reported	12.0 (75)
Varma 2012 [46]	Denmark	Cohort, prospective (20 months)	Hospital	Senior medical consultants Sample size: 2,790 (62%)	Long work hours (>40hrs/week). Self-reported.	Depression: redemption of anti-depressive drug prescriptions	15.0 (94)
Zahrai 2011 [47]	Canada	Cohort, prospective (6 months)	Hospital	Orthopaedic surgery residents Sample size: 16 (not specified)	Night float (n=9): 77.8% did >80h/week at baseline; 71.4% at follow-up). Standard call (n=7): 57.1% did >80h week at baseline; 80% at follow-up.	SF-36, mental health score, self-reported	11.0 (69)

* Follow-up period in brackets for prospective cohort studies; GHQ-12= General Health Questionnaire-12 items; SCLDS= Symptom Checklist Depression Scale; BDI = Beck Depression Inventory II.

Five had been carried out in North America (4 in the United States [37, 38, 40, 44] and 1 in Canada [47]), five in European countries (Denmark,[46] Germany,[42] Sweden,[45] and the United Kingdom [39]) and one study in Japan.[41] Four used a prospective cohort design (follow-up ranging from 6 to 20 months), 2 were case-crossover, 4 cross-sectional, and 1 was a before and after study. Overall, the total number of participants was 14,338 and included medical residents, junior doctors or house officers only working in hospitals, medical specialists or consultants, medical and dental practitioners, general practitioners, hospital physicians and one study included female physicians from a community service. Most studies (n=9) compared different working hour patterns within the same group of physicians, one compared physicians from two different countries,[43] and one included other health care workers as the comparison group.[40]

LWH was defined as more than 48hours/week in most studies, and some did not provide a definition. Two studies considered LWH below 48/week.[43, 45] All the included studies investigated health effects associated with working more than 48h/week, with number of hours ranging between <30 to 110hours/week. One paper studied the effects of an increasing number of hours from 30 to more than 50h/week.[41] Others studied more than 40h/week,[46] 47h/week,[45] above 58h/week [43](two studies) and 6 studies referred to more than 70h/week,[37-40, 44, 47] one of which compared 78-80h/week to 90-110 weekly hours.[44] Information on working hours was self-reported in nine studies, with reasonably good validity only in two studies [37, 38] and reliability in one.[40] In the other two studies it was obtained from established work schedules.[44, 47]

Identified health outcomes that fulfilled the inclusion criteria were percutaneous injuries, motor vehicle accidents, mood disorders and general health. Information was self-reported in all studies using well known validated questionnaires, except three studies that used documented information on motor vehicle crashes,[37] reported percutaneous injuries [39] and antidepressant prescription data as a surrogate of depression.[46]

Summary of findings:

Table 3 shows the findings from each of the 11 studies.

Table 3. Results on the association between long working hours and identified health effects in physicians.

Health outcome	Main results	Adjustment by confounders	Study ID
Percutaneous injuries	Odds ratio (95% CI) for injuries during extended (>20 hours/day) vs. non extended periods (first 12 hours at work): All percutaneous injuries: 1.61 (1.46-1.78); Injuries reported to OH: 1.83 (1.48-2.28); Injuries in the ICU: 1.87 (0.69-5.04); Injuries in the operating room or labour and delivery: 1.77 (1.49-2.09); Injuries in the ICU, non-ICU, or ED: 2.17 (1.56-3.00)	Not adjusted	**Ayas 2006**[37]
Percutaneous injuries	Medical trainees vs. other health care workers: total median working hours: 70 vs. 40 hours (p<.001); median previous working hours at the time of injury: 6.5 vs. 5 hours (p< .001). Odds ratio (95% CI) for self-reported fatigue:> 40h in the past week: 3.5 (2.06-5.92); during more than 5d in past week: 4.2 (4.46-7.15); >12h at work before injury: 8.58 (3.7-19.86). Incidence rate ratio (95%CI) for injuries and fatigue: all workers: 1.40 (1.03-1.90); medical trainees: 2.94 (1.71-5.07); other HCWs: 0.66-1.42) (p 001 for heterogeneity).	Age and sex	**Fisman 2007**[40]
Motor vehicle accidents	Relative risk (95% CI), after extended shift (≥ 24hrs) vs. non extended shift (<24 hours): Crashes = 2.3 (1.6-3.3); Near miss accidents = 5.9 (5.4–6.3)	Age and sex	**Barger 2005**[38]
Motor vehicle accidents	Incidence rates of driving accidents on house visits: Long hours (mean 58.36 hrs/week) = 0.10 (0.31) vs. short hours (mean 38.17 hrs/week) = 0.02 (0.18) vs. p<0.05	Not adjusted	**Kirkcaldy 1997**[42]
Mood disorders, depression	Hazard ratio (95% CI) hours/week intervals and redemption of anti-depressive drug prescription(reference group: 37-40 h/week): 25–36 h: 0.83 (0.24–2.82); 41–44 h: 0.95 (0.50–1.77); 45–49 h: 0.88 (0.43–1.78); 50–54 h: 0.83 (0.32–2.14); 55–59 h: 0.67 (0.15–2.94); >60 h: 0.48 (0.06–3.68). Cox regression analysis of work hours as a continuous variable: HR=0.93 (0.76–1.13)	Age, sex, marital status, medical specialty, decision authority at work, social support at work, quantitative work demands and previous redemption of AD drug prescription	**Varma 2012**[46]
Mood disorders, GHQ-30	Prevalence of cases by working time (hours/week) ≤ 30 hours: 35.7% (41/115); >30 to 40 hours: 39.0% (39/100); >40 to 50 hours: 37.7% (26/69); >50 hours: 56.8% (46/81); *p*=0.0179 Stepwise multiple logistic regression, >50 h/wk. vs. ≤50 h/wk.: parameter 0.635 (p=0.0293)	Marital status, medical facility, position, and night duty.	**Hayasaka 2007**[41]
Mood disorders, GHQ-12 and SCLDS	No association was found between number of hours worked in a week and depression.	Not adjusted	**Firth Cozens 1987**[39]
Mood disorders, SF-36	Scores when night float vs. standard call at baseline and follow-up: Mental health mean score (SD): baseline = 57.33 (22.63) vs. 65.71 (7.61); follow up = 52.00 (15.49) vs. 60.80 (11.45); p=0.72 Mental health component summary (SD), baseline = 34.84 (14.06) vs. 40.21 (7.61) ; follow up = 30.15 (10.71) vs. 42.40 (6.23); p=0.39 Regression analysis: increased number of hours in hospital correlated with significantly lower SF-36 scores in almost all domains.	Not adjusted	**Zahrai 2011**[47]
Mood disorders, SF-36 and BDI	Improvement of levels before and after duty hours limitations from 90-110 to 78-80 hors/week): Mental health SF-36): no statistically significant improvement Individual questions BDI before and after duty hours limitations, only energy level statistically significant.	Not adjusted	**Stamp 2005**[44]

Health outcome	Main results	Adjustment by confounders	Study ID
Mood disorders, SF-36	Odds ratio (95% CI)for working 47 hours per week or more: Impaired mental health, men: 1.59 (0.95-2.66); women: 1.86 (1.03-3.37)	Age and amount of time in practice	**Sundquist 2000**[45]
General health, SF-36	SF-36: Scores when night float vs. standard call at baseline and follow-up: General health mean score (SD): baseline = 62.11 (17.47) vs. 77.57 (24.25); follow up = 56.43 (24.89) vs. 84.20 (16.50); p=0.41 Physical health component summary (SD): baseline = 46.16 (13.15) vs. 52.01 (13.33); follow up = 39.32 (9.80) vs. 56.15 (2.18); p=0.015 Vitality mean score (SD): baseline = 51.67 (14.58) vs. 51.43 (15.74); follow up = 48.57 (14.92) vs. 51.00 (10.84); p=0.20 Regression analysis demonstrated that the increased number of hours spent in hospital correlated with significantly lower general health, physical function, mental health, role emotional, social function and mental component summary scale scores (all p < 0.05)	Not adjusted	**Zahrai 2011**[47]
General health, SF-36	SF-36, BDI: Improvement of levels before and after duty hours limitations from 90-110 to 78-80 hours/week): Physical health: no statistically significant differences.	Not adjusted	**Stamp 2005**[44]
General health, SF-36	SF-36: Odds ratio (95% CI) adjusted for age and time in practice, for working 47 hours per week or more: Impaired general health: men = 1.66 (1.00-2.77); women = 1.59 (1.00-3.17)	Age and amount of time in practice	**Sundquist 2000**[45]
General health, SF-36	Self-rated health: Odds ratio (95% CI) of good self-rated health (logistic regression): Norwegian work time pattern**= 1.35 (1.03-1.77); working in Norway = 4.17 (3.02 - 5.73).	Age, sex and country of work	**Rosta 2011**[43]

**Not working more than 9 hours a day and having more than 60 hours a month on-call); GHQ-12= General Health Questionnaire-12 items; SCLDS= Symptom Checklist Depression Scale; BDI = Beck Depression Inventory II.

Percutaneous injuries: One study of high quality and one of intermediate quality, with non-adjusted analyses, showed consistent results in medical trainees. Ayas et al. found an increased risk of percutaneous injuries associated with working more than 20 hours/day compared to non-extended hour periods, except in intensive care units;[37] Fisman et al. found an association between self-reported fatigue and the number of working hours at the time of injury and a 3-fold increased risk of sharp injuries in medical trainees compared to other health care workers (median of working hours/week before the injury: 70 vs. 40;p<.001).[40]

Motor vehicle accidents: Two studies of high and intermediate quality showed that long working weeks were associated with an at least two-fold increased risk. Barger et al. found adjusted increased risks for car crashes (OR 2·3; 95%CI 1·6-3·3) and near miss accidents (OR 5·9; 95%CI 5·4-6·3) associated with working extended shifts.[38] Kirkcaldy et al. showed that non-adjusted incidence rates of traffic accidents on house visits was five times as much when physicians worked for a mean of 58 hours/week compared to 38 hours/week (p<0,05).[42]

Mood disorders: Six studies of intermediate and high quality found contradictory results for mood disorders. Three of them, all of intermediate quality, provided evidence of an association between LWH and mental health problems. The prevalence of GHQ-30 cases was higher for female physicians working above 50 hours/week than for those working 30 or less hours/week (p<0.05),[41] and an increased adjusted risk of impaired mental health was found for men and women general practitioners who worked for 47 hours per week or more.[45] Zahrai et al. found that increased number of hours spent by residents in hospital correlated significantly with lower mental health and mental component summary scale scores (SF-36).[47] Three other studies, one of which of high quality, did not find an association between depression or mood disorders and LWH. Varma et al. did not find higher adjusted risks of depression associated with increased number of hours in senior medical consultants, even when they worked for more than 60 hours per week.[46] Depression or mood disorders measured with GHQ-12, SCLDS or BDI, all validated tools for mental health, were not associated with increased number of working hours [39] nor improved after decreasing the number of hours from 90-110 to 78-80 hours per week in residents.[44]

General health: Four studies of intermediate quality analysed the association between LWH and general health. Zahrai et al. found that increased number of hours spent by

residents in hospital correlated significantly with poorer general health, physical function and vitality using SF-36.[47] A comparative study of physicians in two different countries found that Norwegians showed higher non-adjusted prevalence of self-reported good health compared to physicians in Germany who worked longer hours.[43] An increased adjusted risk of impaired general health was found for men and women general practitioners who worked for more than 47 hours per week.[45] However, following the implementation of reducing the number of hours from an average of 90-110hours/week to 78-80hours/week did not lead to an overall improvement of residents self-reported physical health.[44]

Synthesis of the evidence:

The levels of evidence synthesis obtained from the analysis of the 11 studies included in this review are shown in table 4.

Table 4. Available scientific evidence on the health problems associated with long working hours in physicians: levels of evidence synthesis.

Health outcome	Degree of evidence*	Magnitude of the association**	Studies ID
Percutaneous injuries	++	++	Ayas 2006[37], Fisman 2007[40]
Motor vehicle accidents	++	++	Barger 2005[38], Kirkcaldy 1997[42]
Mood disorders	+	+/-	Varma 2012[46], Hayasaka 2007[41], Firth Cozens 2001[39], Zahrai 2011[47], Stamp 2005[44], Sundquist 2000[45]
General health	+	+/-	Zahrai 2011[47], Stamp 2005[44], Sundquist 2000[45], Rosta 2011[43]

* Strong evidence (+++): consistent results in more than 2 studies of high quality; Moderate evidence (++): consistent results in two studies of high quality, or one high quality study and one intermediate, or between some studies of intermediate quality; Insufficient evidence (+): identification of only one study or inconsistent results across studies; Evidence of no association (-): consistent results of a non-association in two or more studies.

** Magnitude of the association: it refers to the magnitude of the association point estimate (RR = relative risk, rate ratio or odds ratio): high (+++) if RR > 3; intermediate (++) if RR = 1·5-3; low (+) if RR < 1·5; no association (-); unclear (+/-).

There is moderate evidence of an association between LWH and percutaneous and motor vehicle accidents in physicians. This evidence comes from one study of high quality and another of intermediate quality both for percutaneous injuries, and for motor vehicle accidents and the magnitude of this association could be intermediate.

For mood disorders and general health there is a low or insufficient level of evidence of an association with LWH in physicians. This comes from 6 studies on mood disorders and 4 on general health of high and intermediate methodological quality and inconsistent results among them. No conclusion can be drawn about the magnitude of such associations if existed.

DISCUSSION

This review found moderate scientific evidence for a positive association of intermediate magnitude between exposure to LWH and percutaneous injuries and motor vehicle accidents in physicians. Evidence was assessed as low or insufficient for mood disorders and general health.

To our knowledge, this is the first systematic review on the effects of LWH on physicians' health and safety. The extensive searches were restricted to indexed journals. We systematically identified, selected and assessed the methodological quality of studies by means of independent pairs of reviewers. The quality assessment form and the levels of evidence used in this review were based on CONSORT and STROBE statements and on previous systematic reviews.[33, 35, 36] To give some estimate of the potential magnitude of the effect, we added levels based on the point estimates of measures of association provided by the studies. However, such levels had not been previously established.

Strengths and limitations of the studies

Four studies had a longitudinal prospective design, however only two of them used a robust methodology. Two other studies used a crossover design, which reduces the likelihood of individual variability and confounding, and were considered well designed. Four studies, all rated as intermediate methodological quality, were cross sectional in design and therefore it is difficult to draw causal relationships from them.

Although reverse causality cannot be ruled out from cross-sectional designs, it is unlikely that poorer health determines longer working hours than their healthier peers. One study used a quasi-experimental design but lacked a control group.

Other limitations include the variability, validity and reliability of working hours, which in most studies was self-reported. Although only those papers specifying the number of working hours were accepted, shift and night work might have worked as confounders too, as it is difficult to acknowledge which proportion of the health effects observed was due exclusively to the LWH component and not due, for instance, to stress, fatigue or sleep deprivation. Moreover, the studies did not discriminate between time spent actively working or asleep while on call, though the EWTD considers all hours on call as working time.

Health outcomes were measured mainly based upon self-report, with the exception of documented motor vehicle crashes,[37] reported percutaneous injuries [39] and register-notified anti-depressive drug prescriptions.[45] No studies with other objective measures of health (e.g., mortality) or mental disease (e.g., hospital data) were identified. However, self-reported or perceived health was assessed by validated and widely used instruments, especially when health status (general, mental or physical health) or ill-health symptoms are evaluated. Nonetheless, this raises the possibility that the observed associations might reflect differences in propensity to report health problems when they occur rather than true differences in the risk of worse health status. We cannot rule this out, but the higher propensity for reporting among those who work for longer hours and poorer work schedule patterns would have to relate specifically to the reporting of ill-health rather than non-ill health. It seems unlikely that major differences in propensity to report would extend to a more concrete outcome.

The majority of the studies of this review took into account confounding variables in their analyses, such as sex and age. However, none of the studies analysed the potential effect of modifying factors, such as psychosocial aspects at work, including attitudes, motivation, job requirements, demands and content, organisational climate, social relationships at work, work satisfaction, supportive organisation, or the relative number of physicians and other staff available for patient care. Neither were important aspects outside of work, such as life events or lifestyles, taken into account. The study of Japanese female physicians found significant differences by marital status with higher GHQ scores for those who were married.[41] There are significant differences in the

number of physicians per capita across different countries and this together with the structure and organisation of health services are potential confounders too that were not addressed in the selected studies across countries.

A relevant outcome from this review would be to establish a threshold number of extended hours above which there is a significantly higher risk. There is a lack of evidence for a dose-response relationship and nor does the evidence give any indication for a threshold number of hours that physicians could work and remain safe and healthy.

Finally, the heterogeneity of the included studies did not allow a more quantitative synthesis, and a level of evidence approach was used instead.

Comparison with other studies

Our findings of an association between LWH and injuries and accidents in physicians are consistent with research in other occupational groups. Dembe et al. carried out a longitudinal survey in the US including more than 10,000 workers from a variety of occupations and settings.[16] A strong dose-response effect (adjusted for age, gender, occupation, industry and region) was found between LWH and injuries above 40h/week and 8h/day; jobs with overtime schedules were associated with a 61% higher injury hazard rate compared to jobs without overtime; working at least 12 hours per day was associated with a 37% increased hazard rate, and working at least 60 hours per week was associated with a 23% increased hazard rate. They concluded that LWH might indirectly precipitate workplace injuries through a causal process by inducing fatigue or stress in workers. In our review we found that information about the length of shift work varied across studies, some analysing shifts of up to twenty four hours. Work that included such shifts was more strongly associated with accidents and injuries.[37, 38]

We found insufficient evidence of an association between LWH and mood disorders, other diseases and general health in physicians, despite some scientific evidence suggesting that LWH increase morbidity and mortality in other occupations.[16-23]

Possible explanations for these inconsistencies might be the scarcity of high quality research, and the fact that some of the included papers did not fully address the effect of confounding factors, and the possibility that working as a physician might have a

protective effect by itself, for instance through higher motivation and social recognition relative to other occupations.

Stamp et al., in the only before and after study, did not find significant differences for mood disorder or general health before and after implementing requirements to reduce working schedules to 80h/week in residents.[44] It is debatable, however, whether any conclusions can be drawn from the impact of reducing extremely high working hours (90-110hours/week) to very high working hours (78-80hours/week).

Although we searched for and identified literature describing associations between LWH and increased mortality in doctors,[48, 49] none of those studies specified the number of hours worked, and therefore were not included.

Recommendations for future research

Further research, including well designed prospective and retrospective longitudinal studies, is needed in this field to establish on scientific grounds what number of hours are safe, both to patients and physicians. Also studies on long term health effects should be considered.

This review was aimed to study exclusively possible detrimental health effects of LWH (more than 48 hours per week) to physicians. Other systematic reviews should be conducted addressing issues, such as the impact of working hours on doctor's well-being and quality of life, the quality of junior doctors training, the quality and continuity of care and impact on patient safety. Also, from a staffing perspective, the economic and social burden of sick leave and physicians leaving the profession are important issues that were beyond the scope of this review and that should be considered in future research. Other established criteria [50] that working hours should be 'family friendly', promote gender equality, enhance productivity or facilitate worker choice and influence over working hours, were not addressed. Further research is needed in these other areas.

Conclusions

The findings of this review have confirmed that long working hours are associated with both an increased risk of needlestick injuries and road traffic accidents. The fact that these associations are labelled moderate in this review, should not preclude safety

procedures and measures being assessed and reassessed to ensure maximum physician and patient safety. The reduction of needlestick injuries requires improved use of safe sharp devices, compliance with safe working procedures, training and regular audits of working practices, particularly for junior doctors who are at greatest risk of needlestick injuries, possibly due to inadequate training.[51, 52] The risk of road traffic accidents could be reduced by discouraging driving after very long shifts (more than 16 hours) and where there is sleep deprivation.[53] Transport should be provided by the employer in these situations. This is a pragmatic recommendation which recognises that long working hours have not been eliminated as a result of the EWTD.

This review uncovered the lack of literature on the effect of long working hours on the general health of doctors, a topic that has important health ramifications not only for physicians but for patients too. However, the paucity of available studies, which does not allow for a causal or dose-response relationship to be established, raises further questions about the evidence-base of the current 48 hour limit, especially as this is not enforced strictly and junior doctors, for example in the UK, may opt-out.[54] It may be relevant that the EWTD was not based entirely on science but also took into account, as does all EU legislation, the views of "social partners".[3]

It is likely that the mechanism for these increased health and safety risks is fatigue, and our finding of increased risk of accidents to staff, may also be associated with increased risk of clinical errors. Patient safety was not the purpose of this review but long working hours are well recognised to cause decrement in performance both in health care and other professions, where the performance of staff can be safety critical, such as airline pilots, and professional drivers.[55]

In addition the EWTD has been associated with real concerns about the adequacy of the training of doctors given their reduced exposure to patient care, and the lack of experience of the patient journey, imposed by adherence to the EWTD.[4] This is potentially exacerbated at the same time by the constraints on the overall duration of training imposed by other EU directives.[56] Further reviews of the EWTD, with regard to this unique group of workers, need to take into account other social factors such as the impact on patient care. This is particularly the case when all health care systems are under strain because of burgeoning demands and limited resources.[4]

The findings of this systematic review lead to the suggestion of the following guidelines in the event of non-compliance with the EWTD: 1) long shifts should be avoided to protect both physicians' health and patient safety and no shift should be longer than 16 hours (and then exceptionally);[25] 2) Physicians should be discouraged from driving after long shifts to reduce the risk of motor accidents;[38, 42] 3) rigorous attention should be paid to reducing the risk of sharps injuries;[37, 40] 4) organisational aspects such as workload and job control, as well as the pattern and distribution of working hours, breaks and recovery periods should be carefully taken into account, in order to avoid fatigue and sleep deprivation that could lead to mistakes and accidents;[57, 58] 5) Physicians should work in supportive psychosocial environments, in teams, and with adequate training and supervision as their wellbeing is important also for the health of their patients.[57, 58]

Further research is required to establish any longer term effects such as on mortality and mental health, also the impact of the nature of the work organisation taking into account the psychosocial aspects of the physicians working and non-working lives, and to determine how many hours are safe for physicians to work.

ACKNOWLEDGMENTS

Leena Isotalo, Trial Search Coordinator at the Cochrane Occupational Safety and Health Review Group, designed and conducted the final search of the evidence using Medline and Embase databases.

CONFLICT OF INTERESTS

None. No financial relationships with any organisations that might have an interest in the submitted work, and no other relationships or activities that could appear to have influenced the submitted work.

FUNDING

The Section of Occupational Medicine of the UEMS provided limited support to the work of all authors (MCR, ED, SVP, KS, AS, PR, RH), with the exception of the senior authors (EBM, CS). The expenses of one meeting of the core group (MCR, ED, SVP, EBM, CS) were also funded by the Section. **Role of funding source:** The Section of

Occupational Medicine of the UEMS acted in its role of promoting and contributing to the research in Occupational Medicine.

AUTHORS' CONTRIBUTIONS:

Mari Cruz Rodríguez-Jareño, Evangelia Demou, Sergio Vargas-Prada, Alenka Škerjanc, Kaveh A Sanati, Ewan B. Macdonald and Consol Serra conceived and submitted the design, carried out the acquisition of data, analysis and interpretation of data, drafted the article and revised it.

Pedro G. Reis and Ritva Helimäki-Aro contributed to acquisition of data, analysis and interpretation of data, drafted the article and revised it.

All authors gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

ETHICS COMMITTEE APPROVAL was not sought as no patients were involved in this study.

REFERENCES

1 EUR-Lex. Directive 2003/88/EC of the European Parliament and of the Council of 4 November 2003 concerning certain aspects of the organisation of working time. In: Union OJotE, ed. Brussels: Official Journal of the European Union, 2003.

2 EPHA. 2012 EPHA Briefing on the European Working Time Directive (EWTD) http://www.epha.org/IMG/pdf/EPHA_Briefing_EWTD_Oct2012_FINAL.pdf; Access Year 2012.

3 European Commission. 2013 Working Conditions-Working Time Directive. http://ec.europa.eu/social/main.jsp?catId=706&langId=en&intPageId=205; Access Date December 2013.

4 House J. Calling time on doctors' working hours. *Lancet* 2009;**373**:2011-12.

5 The Lancet. Doctors' training and the European Working Time Directive. *Lancet* 2010;**375**:2121.

6 Leff D, Aziz O, Darzi A. Trucks, planes, and scalpels - Is there an evidence-based approach to surgeons' working hours? *Arch Surg* 2007;**142**:817-20.

7 Parthasarathy S. Sleep and the medical profession. *Curr Opin Pulm Med* 2005;**11**:507-12.

8 Dr. Foster Hospital Guide. 2011 www.drfosterhealth.co.uk; Access Date February 2011.

9 UEMS. 2012 The Union of European Medical Specialists. www.uems.net; Acess Year 2012.

10 Macdonald EB, Ritchie KA, Murray KJ, Gilmour WH. Requirements for occupational medicine training in Europe: a Delphi study. *Occup Environ Med* 2000;**57**:98-105.

11 Macdonald E, Baransky B, Wilford J. Occupational Medicine in Europe: Scope and Cometencies. In: Health WECfEa, ed. Health, Environment and Safety in Enterprises Series n. 3. Bilthoven, The Netherlands, 2000.

12 Ballester M, Cordell N, Rodriguez Jareno MC, Serra C, Med USO. A European survey of professional bodies representing occupational medicine specialists. *Occup Med (Oxf)* 2012;**62**:366-70.

13 Cashman C, Slovak A. The occupational medicine agenda: Routes and standards of specialization in occupational medicine in europe. *Occup Med (Oxf)* 2005;**55**:308-11.

14 De Schryver A, Claesen B, Meheus A, van Sprundel M, Francois G. European survey of hepatitis B vaccination policies for healthcare workers. *Eur J Public Health* 2011;**21**:338-43.

15 Bannai A, Tamakoshi A. The association between long working hours and health: A systematic review of epidemiological evidence. *Scand J Work Environ Health* 2013.

16 Dembe AE, Erickson JB, Delbos RG, Banks SM. The impact of overtime and long work hours on occupational injuries and illnesses: new evidence from the United States. *Occup Environ Med* 2005;**62**:588-97.

17 Johnson JV, Lipscomb J. Long working hours, occupational health and the changing nature of work organization. *Am J Ind Med* 2006;**49**:921-9.

18 Kivimaeki M, Batty GD, Hamer M, et al. Using Additional Information on Working Hours to Predict Coronary Heart Disease. *Ann Intern Med* 2011;**154**.

19 Liu Y, Tanaka H, Fukuoka Heart Study G. Overtime work, insufficient sleep, and risk of non-fatal acute myocardial infarction in Japanese men. *Occup Environ Med* 2002;**59**:447-51.

20 Sokejima S, Kagamimori S. Working hours as a risk factor for acute myocardial infarction in Japan: case-control study. *BMJ* 1998;**317**:775-80.

21 Spurgeon A, Harrington JM, Cooper CL. Health and safety problems associated with long working hours: A review of the current position. *Occup Environ Med* 1997;**54**:367-75.

22 Virtanen M, Ferrie JE, Singh-Manoux A, et al. Long working hours and symptoms of anxiety and depression: a 5-year follow-up of the Whitehall II study. *Psychol Med* 2011;**41**:2485-94.

23 Virtanen M, Stansfeld SA, Fuhrer R, Ferrie JE, Kivimaki M. Overtime work as a predictor of major depressive episode: a 5-year follow-up of the Whitehall II study. *PloS One* 2012;**7**:e30719.

24 Gopal R, Glasheen JJ, Miyoshi TJ, Prochazka AV. Burnout and internal medicine resident work-hour restrictions. *Arch Intern Med* 2005;**165**:2595-600.

25 Reed DA, Fletcher KE, Arora VM. Systematic Review: Association of Shift Length, Protected Sleep Time, and Night Float With Patient Care, Residents' Health, and Education. *Ann Intern Med* 2010;**153**:829-42.

26 Sargent MC, Sotile W, Sotile MO, Rubash H, Barrack RL. Quality of Life During Orthopedic Training and Academic Practice Part 1: Orthopedic Surgery Residents and Faculty. *J Bone Joint Surg Am* 2009;**91A**:2395-405.

27 Fletcher KE, Underwood W, Davis SQ, Mangrulkar RS, McMahon LF, Saint S. Effects of work hour reduction on residents' lives - A systematic review. *JAMA* 2005;**294**:1088-100.

28 Black C, Frost D. *Health at work - an independent review on sickness absence*. London: The Stationary Office, 2011.

29 Waddell G, Burton K. *Is work good for your health and wellbeing?* London: The Stationary Office, 2006.

30 Butterworth P, Leach LS, Strazdins L, Olesen SC, Rodgers B, Broom DH. The psychosocial quality of work determines whether employment has benefits for mental health: results from a longitudinal national household panel survey. *Occup Environ Med* 2011;**68**:806-12.

31 Lockley SW, Landrigan CP, Barger LK, Czeisler CA. When policy meets physiology - The challenge of reducing resident work hours. *Clin Orthop* 2006:116-27.

32 ICD-10. 2012 International Classification of Diseases. http://www.who.int/classifications/icd/en/; Access Year 2012.

33 van Uffelen JGZ, Wong J, Chau JY, et al. Occupational Sitting and Health Risks A Systematic Review. *Am J Prev Med* 2010;**39**:379-88.

34 Kirsling RA, Kochar MS, Chan CH. An evaluation of mood states among 1st-year residents. *Psychol Rep* 1989;**65**:355-66.

35 Bernard B. A critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity and low back. In: NIOSH, ed. Publication No. 9741. Cincinnati: National Institute for Occupational Safety and Health, 1997.

36 Steenstra IA, Verbeek JH, Heymans MW, Bongers PM. Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature. *Occup Environ Med* 2005;**62**:851-60.

37 Ayas NT, Barger LK, Cade BE, et al. Extended work duration and the risk of self-reported percutaneous injuries in interns. *JAMA* 2006;**296**:1055-62.

38 Barger LK, Cade BE, Ayas NT, et al. Extended work shifts and the risk of motor vehicle crashes among interns. *N Engl J Med* 2005;**352**:125-34.

39 Firth-Cozens J. Emotional Distress in Junior House Officers. *BMJ* 1987;**295**:533-36.

40 Fisman DN, Harris AD, Rubin M, Sorock GS, Mittleman MA. Fatigue increases the risk of injury from sharp devices in medical trainees: Results from a case-crossover study. *Infect Control HospEpidemiology* 2007;**28**:10-17.

41 Hayasaka Y, Nakamura K, Yamamoto M, Sasaki S. Work environment and mental health status assessed by the general health questionnaire in female Japanese doctors. *Ind Health* 2007;**45**:781-86.

42 Kirkcaldy BD, Trimpop R, Cooper CL. Working hours, job stress, work satisfaction, and accident rates among medical practitioners and allied personnel. *Int J Stress Manag* 1997;**4**:79-87.

43 Rosta J, Aasland OG. Work Hours and Self rated Health of Hospital Doctors in Norway and Germany. A comparative study on national samples. *BMC Health Serv Res* 2011;**11**.

44 Stamp T, Termuhlen P, Miller S, et al. Before and after resident work hour limitations: an objective assessment of the well-being of surgical residents. *Curr Surg* 2005;**62**:117-21.

45 Sundquist J, Johansson SE. High demand, low control, and impaired general health: working conditions in a sample of Swedish general practitioners. *Scand J Pub Health* 2000;**28**:123-31.

46 Varma A, Marott J, Stoltenberg C, Wieclaw J, Kolstad H, Bonde J. With long hours of work, might depression then lurk? A nationwide prospective follow-up study among Danish senior medical consultants. *Scand J Work Environ Health* 2012;**38**:418-26.

47 Zahrai A, Chahal J, Stojimirovic D, Schemitsch EH, Yee A, Kraemer W. Quality of life and educational benefit among orthopedic surgery residents: a prospective, multicentre comparison of the night float and the standard call systems. *Can J Surg* 2011;**54**:25-32.

48 Lindfors PM, Nurmi KE, Meretoja OA, et al. On-call stress among Finnish anaesthetists. *Anaesthesia* 2006;**61**:859-64.

49 Šelb J, Albreht T. Mortality rates of medical doctors in Slovenia in 1985 to 1999. *Zdravniški vestnik* 2000:I47-I48.

50 Lee S, McCann D, Messenger J. Working Time Around the World: Trends in working hours, laws and policies in a global comparative perspective. Geneva, Switzerland: International Labour Office, 2007.

51 Elder A, Paterson C. Sharps injuries in UK health care: a review of injury rates, viral transmission and potential efficacy of safety devices. *Occup Med-Oxford* 2006;**56**:566-74.

52 Naghavi SHR, Sanati KA. Needlestick injuries: Does left-handedness matter? *Am J Infect Control* 2009;**37**:341-41.

53 Connor J, Norton R, Ameratunga S, et al. Driver sleepiness and risk of serious injury to car occupants: population based case control study. *BMJ* 2002;**324**:1125-28A.

54 BMA. 2013 European Working Time Directive, http://bma.org.uk/practical-support-at-work/ewtd,. Access Date December 2013.

55 Wilson AM, Weston G. Application of Airline Pilots Hours to Junior Doctors. *BMJ* 1989;**299**:779-81.

56 The European Specialist Medical Qualifications Order 1995. 1995 Statutory Instruments, 1995 No. 3208, MEDICAL PROFESSION, The European Specialist Medical Qualifications Order 1995, http://legislation.data.gov.uk/uksi/1995/3208/made/data.htm?wrap=true,. Access Date December 2013.

57 Boorman S. NHS Health and Well-being. The Boorman Review. Leeds, 2009.

58 Williams S, Michie S, Pattani S. Improving the health of the NHS workforce. Report of the partnership on the health of the NHS Workforce. London: The Nuffield Trust, 1998.

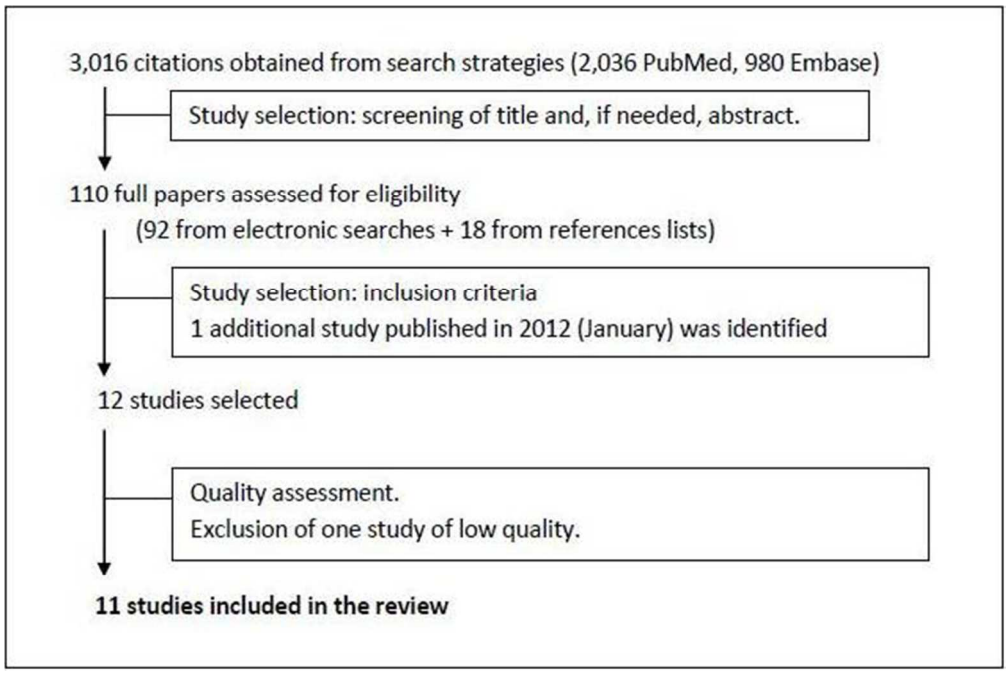


Figure 1.Results of the search strategy, using search engines on PubMed and EMBASE (December 2011) and screening of references lists of identified full papers, study selection and quality assessment.
168x112mm (96 x 96 DPI)

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6-7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	6
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	8-9
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	8-9
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	8-9
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8-9
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	9
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	8-9

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	7
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	10-11
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	12
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	17
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	9-18
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	18
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	18-20
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	20-22
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	23-24

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

BMJ Open

European Working Time Directive and doctors' health. A Systematic review of the available epidemiological evidence.

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-004916.R1
Article Type:	Research
Date Submitted by the Author:	24-May-2014
Complete List of Authors:	Rodriguez-Jareno, Maria-Cruz; Universitat de Girona, 2. Department of Medical Sciences Demou, Evangelia; University of Glasgow, Institute of Health and Wellbeing Vargas, Sergio; Universitat Pompeu Fabra, Centre for Research in Occupational Health Sanati, Kaveh; University of Glasgow, Institute of Health and Wellbeing Skerjanc, Alenka; University Medical Centre, 6. Clinical Institute of Occupational, Traffic and Sports Medicine Reis, Pedro; 7. College of Occupational Medicine, Portuguese Medical Association Helimaki-Aro, Ritva; 8. Helsinki City Occupational Health Centre, Macdonald, Ewan; University of Glasgow, Institute of Health and Wellbeing Serra, Consol; Universitat Pompeu Fabra, Centre for Research in Occupational Health
Primary Subject Heading:	Health policy
Secondary Subject Heading:	Medical management, Research methods, Health services research, Health policy, Evidence based practice
Keywords:	EPIDEMIOLOGY, GENERAL MEDICINE (see Internal Medicine), HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

European Working Time Directive and doctors' health. A Systematic review of the available epidemiological evidence.

Running title: Long working hours and doctors' health

***Corresponding author**

Dr. Evangelia Demou
Healthy Working Lives Group
Institute of Health and Wellbeing
University of Glasgow
Glasgow, G12 8RZ
Tel: +44 141 330 3559
E-mail: evangelia.demou@glasgow.ac.uk

Maria Cruz Rodriguez-Jareño[1,2], Evangelia Demou[3,*], Sergio Vargas-Prada[4], Kaveh A Sanati[3,5], Alenka Škerjanc[6], Pedro G. Reis[7], Ritva Helimäki-Aro[8], Ewan B. Macdonald[3], Consol Serra[1,4,9,10], on behalf of the UEMS Section of Occupational Medicine[11].

1. Catalan and Spanish Societies of Occupational Medicine, Barcelona, Spain.
2. Department of Medical Sciences. School of Medicine. Universitat de Girona. Girona. Spain
3. Institute of Health and Wellbeing, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom
4. CiSAL - Centre for Research in Occupational Health, Universitat Pompeu Fabra, Barcelona, Spain.
5. Occupational Health Department, Epsom & St Helier University Hospitals NHS Trust, Surrey, United Kingdom
6. Clinical Institute of Occupational, Traffic and Sports Medicine, University Medical Centre, Ljubljana, Slovenia.
7. College of Occupational Medicine, Portuguese Medical Association, Lisbon, Portugal.
8. Helsinki City Occupational Health Centre, Helsinki, Finland.
9. Occupational Health Service, Parc de Salut MAR, Barcelona, Spain.
10. CIBER of Epidemiology and Public Health. Spain.
11. UEMS Section of Occupational Medicine. http://www.uems-occupationalmedicine.org

Word Count: 4,794

ABSTRACT

Objective: To summarise the available scientific evidence on the health effects of exposure to working beyond the limit number of hours established by the European Working Time Directive (EWTD) on physicians.

Design: A systematic literature search was conducted in PubMed and Embase. Study selection, quality appraisal and data extraction were carried out by independent pairs of researchers using pre-established criteria.

Setting: Physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.

Participants: The total number of participants was 14,338.

Primary and secondary outcome measures: Health effects classified under the International Classification of Diseases (ICD-10).

Results: Over 3,000 citations and 110 full articles were reviewed. From these, eleven studies of high or intermediate quality carried out in North America, Europe and Japan met the inclusion criteria. Six studies included medical residents, junior doctors or house officers and the five others included medical specialists or consultants, medical, dental and general practitioners and hospital physicians. Evidence of an association was found between percutaneous injuries and road traffic accidents with extended hours/day or very long working weeks (LWH). The evidence was insufficient for mood disorders and general health. No studies on other health outcomes were identified.

Conclusions: LWH could increase the risk of percutaneous injuries and road traffic accidents, and possibly other incidents at work through the same pathway. While associations are clear, the existing evidence does not allow for an established causal or "dose-response" relationship between LWH and incidents at work, or for a threshold number of extended hours above which there is a significantly higher risk and the hours physicians could work and remain safe and healthy. Policy makers should consider safety issues when working on relaxing EWTD for doctors.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- The first systematic review, to our knowledge, on the effects of Long Working Hours (LWH) on physicians' health and safety.
- A systematic literature search conducted in PubMed and Embase with over 3,000 citations and 110 full articles reviewed.
- Eleven studies of high or intermediate quality carried out in North America, Europe and Japan, involving physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.
- The findings of this review confirmed that long working hours are associated with both an increased risk of needlestick injuries and road traffic accidents. Evidence was assessed as low or insufficient for mood disorders and general health.
- This review uncovered the lack of literature on the effect of long working hours on the general health of doctors, a topic that has important health ramifications not only for physicians but for patients too.

KEY WORDS: Working time, "Physicians"[MeSH], "Morbidity"[MeSH], "Occupational Injuries"[MeSH], Systematic review.

INTRODUCTION

The European Working Time Directive (EWTD),[1] applicable to all occupations across the EU, requires a maximum working week of 48 hours and establishes rest periods.[2, 3] Since it came into force in healthcare in 2009, it has been associated with concerns about the provision of health services including continuity of care, lower staffing levels, introduction of shift working, a reduction in training time and the adequate supervision of junior doctors.[4, 5] The large inter-country variations in physicians working hours as well as the mandatory controls over work hours across occupations raises further questions as to the basis of selecting and setting these limits and restrictions.[6] There is evidence of variable compliance to the EWTD within health care across Europe.[4, 5] The medical profession is also increasingly feminised with potential implications for women of childbearing age, which may coincide with the time in the profession where long working hours (LWH) are more likely to happen (i.e. during residency). Constraint on public-sector finances and increasing health care demands, have stimulated the debate on physicians' working hours,[7] with some evidence showing that the quality of health care is positively correlated with the number of physicians available to deliver that care.[8]

The Union of European Medical Specialists (UEMS) was created in 1958 to represent medical specialists in the EU, promote a high standard of training and practice at European level and facilitate the free movement of physicians across European countries.[9] The Occupational Medicine section was created in 1997 and is involved in issues related to specialist training, professional practice and research;[10] elaborating and publishing reference documents,[11] conducting and contributing to surveys,[12-14] and working in partnership with other key European groups.

The trigger for this project was a request by the Council of the UEMS to its Occupational Medicine Section for a position statement on the EWTD and possible health consequences to physicians of a potential removal of this professional group from the current EWTD 48 hours per week limit.[15-17] An evidence based approach was taken by the Section and a systematic review of the literature was undertaken.[15]

LWH have been associated with several adverse effects on workers' health, such as hypertension, cardiovascular disease, stress, depression, musculoskeletal disorders, chronic infections, diabetes, general health complaints and all-cause mortality in a variety of occupational groups.[18-26] In physicians, there is some evidence suggesting

that LWH can result in impacts such as stress, depression, burnout, injuries, fatigue and sleep deprivation,[27-29] and overwork has been considered one of the most stressful features of physicians' work.[27] Some evidence exists of the beneficial effect of reducing the number of working hours on the quality of life of medical residents [30] and on burnout.[27]

However, the relationship between work and health is complex. There is a substantial body of evidence showing that worklessness is associated with poorer health, whereas work is generally good for health and well-being,[31, 32] provided that the work environment is reasonably acceptable and supportive.[33]

Occupational and non-occupational exposures may play an important role such as the pattern and distribution of working hours, breaks, and recovery periods. Also, psychosocial and organisational factors, such as workload, job control, managers and peer support, training opportunities, individual characteristics and attitudes may be important.[6, 34]

Finally, physicians are a highly qualified, devoted and motivated professional group. Despite other existing reviews of a more general occupational approach, a review focused on physicians is warranted.

The aim of this project was to systematically review whether long-working hours (LWH) -i.e. defined as more than the 48 hour per week limit imposed by the EWTD- are associated with health effects –i.e. classified under the International Classification of Diseases (ICD-10) [35]- in physicians, and to examine what these associations are. This review does not include the potential impact of LWH on patient care or physician training.

METHODS

The systematic review sought to respond to the following research question: are LWH associated with health effects in physicians? For the purpose of this review, exposure to LWH was defined as working for more than 48 hours per week.[1] Health outcomes included any disease as defined in the ICD-10 [35] and work related injuries. If self-assessed tools on health status were used, only those validated were considered. Burnout, stress and outcomes such as individual symptoms, signs or biological markers (blood pressure, ECG, etc.) were not considered. Physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.

Study identification

Electronic searches were carried out using PubMed and EMBASE as search engines (December 2011). Our search strategy was similar in both databases and consisted of a combination of the following keywords and Mesh terms: night shift, morning shift, evening shift, afternoon shift, shift work, rotating shift, shift combination, shift duration or length, shift system, clockwise rotation, shift roster, extended shifts, night work, evening work, work schedule, work hours, starting time, early start, irregular working hours, direction of rotation, overwork, extended hours, shift rota, workload, work schedule tolerance, sleep deprivation, sleep disorders, chronobiology disorders, circadian rhythm, psychomotor performance, circadian disruption, vigilance, alertness, wakefulness, drowsiness, fatigue, insomnia, hypersomnolence, dyssomnia, eveningness, morningness, neurocognitive performance, concentration difficulties, arousal, health, morbidity, mortality, disease, illness, stress, strain, distress, accident, injur*, death, suicid*, education, medical, physician, medical staff, hospital, doctor, surgeon, house officers, medical school, surgery, surgical.

In addition, the references lists of articles selected for inclusion were carefully reviewed to identify additional studies.

Study selection

Studies were included if they measured the association between the exposure to LWH and health effects in physicians, used an observational epidemiological (i.e. cohort, case-control, cross-sectional) or experimental design. Articles relating to on-call, night or shift work, but with unknown exposure or exposed to less than 48 working hours per week were excluded. They were also excluded if the working groups were other than physicians or addressed only other exposures or outcomes (i.e. patient safety, fatigue, sleep deprivation, social or family disruption). No limitation was set for languages.

A total of 2,036 citations were obtained from PubMed and 980 non-duplicated citations using Embase, yielding a total of 3,016 citations that were all screened by independent pairs of reviewers. All titles and, where necessary, abstracts were reviewed. Ninety two potentially suitable publications were identified from the electronic searches, and 18 further studies were identified from the references lists, yielding a total of 110 studies published in English, French, German, Italian, Slovenian and Spanish. For all of them the full text was obtained and reviewed by independent pairs of reviewers. Disagreements within pairs were resolved by discussion and, where necessary, by a third reviewer who made the final decision. All authors participated as reviewers for screening the citations and full papers using well defined and pre-established criteria. Ninety nine articles were excluded at this stage. One further study was identified that was published after the search period. Finally, 12 papers were considered for quality assessment (Figure 1).

The methodological quality of the 12 studies was assessed by independent pairs of reviewers. A standardized 16 point scale based on CONSORT and STROBE statements and adapted from a previous systematic review [36] was used. It includes 16 items grouped into 6 areas: 1) objectives, 2) study design, 3) target population and sample, 4) variables, 5) data sources, collection and measurement, and 6) statistical methods. Each item was rated as 1 (the requirement was met), 0.5 (the requirement was partially met) or 0 (the requirement was not or unclearly met). Disagreements within pairs of reviewers were resolved by consensus or, where necessary, by a third reviewer. For each of the 12 studies, a final score based on the sum of all items was assigned and the percentage was calculated based on a maximum score of 16. Study quality was rated as low, moderate or high if it scored less than 60%, between 60% and 79.9%, and 80% or more of the maximum score, respectively.

One article of low quality was excluded [37] and a total of 11 studies of intermediate or high methodological quality were included for the purpose of this review (figure 1).

Data extraction

Selected information was extracted from each paper, including publication year, country of origin, study design, setting, study population, sample size, response rate, measure of working hours, definition used for LWH, health outcomes and their measurement, and main results on point risk estimates or frequencies of health outcomes, and their corresponding 95% confidence interval (95%CI), and whether the analyses were adjusted for potential confounders.

Evidence synthesis

To summarise the results on the relationship between LWH and health effects, levels of evidence synthesis was performed. This was based on the methodological quality, study design and the consistency of the study outcomes. The following criteria were based on two previous relevant systematic reviews.[38, 39]

Strong evidence: consistent results in more than two studies of high quality.

Moderate evidence: consistent results in one high quality study and one intermediate, or between some studies of intermediate quality.

Insufficient evidence: identification of only one study or inconsistent results across studies;

Evidence of no association: consistent results of a non-association in two or more studies.

A measure of the possible magnitude of the association was attempted using the following criteria based on the association point estimate (RR = relative risk, rate ratio or odds ratio) when available or otherwise the comparison between two frequencies was used:

- high: RR > 3·00
- intermediate: RR = 1·50-3·00
- low: RR = 1·01-1·49
- no association: RR=1·00

- unclear

RESULTS

We identified 11 studies related to LWH and health effects in physicians that were eligible for inclusion. The outcome of the methodological quality assessment is given in table 1.

Methodological quality was appraised as high for 3 of the 11 studies, with scores ranging from 84 to 94%. The other 8 studies were considered as intermediate quality as their scores ranged from 63 to 78%. The majority of all included studies received positive scores on items describing the study objectives, design and the study population (items from areas 1, 2 and 3), as well as the description of exposure and outcome variables (item 4). However, although data sources and collection were well described, reliability and validity of exposure to LWH and health outcome measures had low scores (items 5b and 5c). The statistical methods were in general appropriate, however the confounding variables were not measured in several studies and, as can be seen, the studies did not describe how they addressed missing data (items 6a and 6b).

Study characteristics:

The characteristics of the included studies are described in table 2.

Table 1. Methodological quality appraisal of the included studies

Study ID			Ayas 2006 [40]	Barger 2005 [41]	Firth-Cozens 1987 [42]	Fisman 2007 [43]	Hayasaka 2007 [44]	Kirkcaldy 1997 [45]	Rosta 2011 [46]	Stamp 2005 [47]	Sundquist 2000 [48]	Varma 2012 [49]	Zahrai 2011 [50]
1	Objectives	Are the objectives or hypotheses of the research described in the paper?	1	1	1	1	1	1	1	1	1	1	1
2	Study design	Is the study design presented?	1	1	1	1	1	1	1	1	1	1	1
3a	Target population	Do the authors describe the target population they wanted to research?	1	1	1	1	1	1	1	1	1	1	1
3b	Sample	Was a random sample of the target population taken? AND was the response rate 60% or more?	0.5	0.5	1	0.5	0.5	0.5	0.5	1	1	1	0.5
3c	Sample	Is participant selection described?	1	1	1	1	0.5	1	0	1	1	1	1
3d	Sample	Is participant recruitment described, or referred to?	1	1	1	1	1	1	1	1	1	1	0
3e	Sample	Are the inclusion and/or exclusion criteria stated?	1	1	1	1	0.5	0.5	1	1	1	1	0.5
3f	Sample	Is the study sample described? (minimum description: sample size, gender, age and occupation)	1	1	0	1	1	1	0	0	0	1	1
3g	Sample	Is the number of participants at each stage of the study reported?	1	1	1	0.5	1	0.5	1	1	1	1	1
4	Variables	Are the measures of long working hours and the health outcome described?	1	1	0.5	1	1	1	1	1	1	1	1
5a	Data sources, collection	Do authors describe the source of their data (e.g., official registry, health survey) AND how the data were collected?	1	1	1	1	1	1	1	1	1	1	1
5b	Measurement	Was reliability of the measure(s) of long working hours mentioned or referred to?	0	0	0	1	0	0	0	1	0	1	0
5c	Measurement	Was the validity of the measure(s) of long working hours mentioned or referred to?	1	1	0	0	0	0	0	0	0	0	0
5d	Measurement	Were health outcomes assessed by objective measures or validated self-reporting instruments?	0	1	1	1	1	0	1	1	1	1	1

6a	Statistical methods	Were appropriate statistical methods used and described, including those for addressing confounders?	0.5	1	0.5	1	0.5	0.5	1	0.5	1	1	0.5
6b	Statistical methods	Were the numbers/% of participants with missing data for long working hours and the health outcome indicated AND If more than 20% of data in the primary analyses were missing, were methods used to address missing data?	0	0	0	0.5	0	0	0	0	0	1	0.5
Total score** (%)			12.0 (75)	13.5 (84)	11.0 (69)	13.5 (84)	11.0 (69)	10.0 (63)	10.5 (66)	12.5 (78)	12.0 (75)	15.0 (94)	11.0 (69)
Quality rate***			interm	high	interm	high	interm	interm	interm	interm	interm	high	interm

** Maximum score = 16; *** Quality rate (%): low=<60; intermediate (interm)=60-79; high=80-100.

Table 2. Characteristics of the included studies.

Study ID	Country Of Study Population	Design*	Setting	Participants and sample size (response rate)	Working hours	Health outcomes	Quality score (%)
Ayas 2006 [40]	United States	Cohort, prospective (1 year) / Case-crossover	Hospital	Interns in postgraduate residency programs Sample size: 2,737 (80%)	Mean (SD) hrs worked/month: 249.8 (75.3). Self-reported, monthly survey. Strong correlation (Pearson *r*=0.76; *p* 0.001) with hours worked 244 (69.3) from work diaries of randomly selected subset of 192 interns. Extended periods (20 or more consecutive hours) vs. non extended periods (12 hrs or less consecutive hours).	Self-reported percutaneous injuries	12.0 (75)
Barger 2005 [41]	United States	Cohort, prospective (1 year)	Hospital	Residents first postgraduate year (interns). Sample size: 2,737 (80%)	Mean (SD) hrs worked/week 70.7 (26.0). Extended shift (≥ 24hrs) vs. non extended shift (<24 hours). Self-reported, validated.	Self-reported and documented motor vehicle crashes.	13,5 (84)
Firth Cozens 1987 [42]	United Kingdom	Cohort, prospective (1 year)	Hospital	Junior house officers Sample size: 170 (72%)	Mean number of hours per week = 90.6 hours (include on call). Self-reported.	GHQ-12 (case: score ≥2) and SCLDS, self-reported	11.0 (69)
Fisman 2007 [43]	United States and Canada	Case-Crossover	Hospital	Medical trainees vs. other HCW. Sample size: 109 vs. 241 (46%)	Median number of hours per week: medical trainees = 70; other HCW = 40 (p<0.001). Self-reported, high reliability.	Reported percutaneous injuries to employee health care service.	13.5 (84)
Hayasaka 2007 [44]	Japan	Cross-sectional	Hospital, clinics, other	Female physicians Sample size: 367 (63%)	Comparison of increasing number of hrs/week from ≤ 30 to >50, self-reported.	GHQ-30 (case: score ≥8), self-reported	11.0 (69)
Kirkcaldy 1997[45]	Germany	Cross-sectional	Not specified	Medical and dental practitioners. Sample size: 2,500 (not specified)	Long hours: mean 58.36 hrs/week SD 9.16 Short hours: mean 38.17 hrs/week SD 7.72 Self-reported.	Self-reported traffic accidents	10.0 (63)
Rosta 2011[46]	Germany and Norway	Cross-sectional	Hospital	Hospital physicians Participants: 1,917 (58%) in Germany and 1,072 (65%) in Norway answered the questionnaire. From survey responders, 1,260 (65.7%) Germans and 562 (52.4) Norwegians were included in the analysis after applying further inclusion criteria.	German vs. Norwegian physicians (%): Hours per day >9 = 58.8 vs. 26.7 60 hours on-call per month = 63.4 vs. 18.3 Self-reported.	Validated questionnaire on self-rated health	10.5 (66)
Stamp 2005 [47]	United States	Before-and-after	Hospital	Residents of general surgery Sample size: 28 (97%)	Changes of work patterns, after reduction from 90-110 to 78-80 work hrs/week.	SF-36, mental health; BDI, depression; self-reported	12.5 (78)

Sundquist 2000 [48]	Sweden	Cross-sectional	Primary care	General practitioners Sample size: 1,004 (72%)	Overtime defined as working at least 47 hrs/week. Self-reported.	Swedish SF-36, impaired mental health, self-reported	12.0 (75)
Varma 2012 [49]	Denmark	Cohort, prospective (20 months)	Hospital	Senior medical consultants Sample size: 2,790 (62%)	Long work hours (>40hrs/week). Self-reported.	Depression: redemption of anti-depressive drug prescriptions	15.0 (94)
Zahrai 2011 [50]	Canada	Cohort, prospective (6 months)	Hospital	Orthopaedic surgery residents Sample size: 16 (not specified)	Night float (n=9): 77.8% did >80h/week at baseline; 71.4% at follow-up). Standard call (n=7): 57.1% did >80h week at baseline; 80% at follow-up.	SF-36, mental health score, self-reported	11.0 (69)

* Follow-up period in brackets for prospective cohort studies; GHQ-12= General Health Questionnaire-12 items; SCLDS= Symptom Checklist Depression Scale; BDI = Beck Depression Inventory II.

Five had been carried out in North America (4 in the United States [40, 41, 43, 47] and 1 in Canada [50]), five in European countries (Denmark,[49] Germany,[45] Norway and Germany [43]; Sweden,[48] and the United Kingdom [42]) and one study in Japan.[44] Four used a prospective cohort design (follow-up ranging from 6 to 20 months), 2 were case-crossover, 4 cross-sectional, and 1 was a before and after study. Overall, the total number of participants was 14,338 and included medical residents, junior doctors or house officers only working in hospitals, medical specialists or consultants, medical and dental practitioners, general practitioners, hospital physicians and one study included female physicians from a community service. Most studies (n=9) compared different working hour patterns within the same group of physicians, one compared physicians from two different countries,[46] and one included other health care workers as the comparison group.[43]

LWH was defined as more than 48hours/week in most studies, and some did not provide a definition. Two studies considered LWH below 48/week.[46, 48] All the included studies investigated health effects associated with working more than 48h/week, with number of hours ranging between <30 to 110hours/week. One paper studied the effects of an increasing number of hours from 30 to more than 50h/week.[44] Others studied more than 40h/week,[49] 47h/week,[48] above 58h/week [46](two studies) and 6 studies referred to more than 70h/week,[40-43, 47, 50] one of which compared 78-80h/week to 90-110 weekly hours.[47] Information on working hours was self-reported in nine studies, with reasonably good validity only in two studies [40, 41] and reliability in one.[43] In the other two studies it was obtained from established work schedules.[47, 50]

Identified health outcomes that fulfilled the inclusion criteria were percutaneous injuries, motor vehicle accidents, mood disorders and general health. Information was self-reported in all studies using well known validated questionnaires, except three studies that used documented information on motor vehicle crashes,[40] reported percutaneous injuries [42] and antidepressant prescription data as a surrogate of depression.[49]

Summary of findings:

Table 3 shows the findings from each of the 11 studies.

Table 3. Results on the association between long working hours and identified health effects in physicians.

Health outcome	Main results	Adjustment by confounders	Study ID
Percutaneous injuries	Odds ratio (95% CI) for injuries during extended (>20 hours/day) vs. non extended periods (first 12 hours at work): All percutaneous injuries: 1.61 (1.46-1.78); Injuries reported to OH: 1.83 (1.48-2.28); Injuries in the ICU: 1.87 (0.69-5.04); Injuries in the operating room or labour and delivery: 1.77 (1.49-2.09); Injuries in the ICU, non-ICU, or ED: 2.17 (1.56-3.00)	Time of day and circadian influences	**Ayas 2006**[40]
Percutaneous injuries	Medical trainees vs. other health care workers: total median working hours: 70 vs. 40 hours (p<.001); median previous working hours at the time of injury: 6.5 vs. 5 hours (p< .001). Odds ratio (95% CI) for self-reported fatigue: work > 40h in the past week: 3.50 (2.06-5.92); work during more than 5d in past week: 4.20 (4.46-7.15); >12h at work before injury: 8.58 (3.7-19.86). Incidence rate ratio (95%CI) for association between fatigue and injuries: all workers: 1.40 (1.03-1.90); medical trainees: 2.94 (1.71-5.07); other HCWs: 0.97 (0.66-1.42) (p=0.001 for heterogeneity).	Age and sex	**Fisman 2007**[43]
Motor vehicle accidents	Odds ratios (95% CI), after extended shift (≥ 24hrs) vs. non extended shift (<24 hours): Crashes = 2.3 (1.6-3.3); Near miss accidents = 5.9 (5.4–6.3)	Age and sex	**Barger 2005**[41]
Motor vehicle accidents	Incidence rates of driving accidents on house visits: Long hours (mean 58.36 hrs/week) = 0.10 (0.31) vs. short hours (mean 38.17 hrs/week) = 0.02 (0.18). p<0.05	Not adjusted	**Kirkcaldy 1997**[45]
Mood disorders, depression	Hazard ratio (95% CI) hours/week intervals and redemption of anti-depressive drug prescription(reference group: 37-40 h/week): 25–36 h: 0.83 (0.24–2.82); 41–44 h: 0.95 (0.50–1.77); 45–49 h: 0.88 (0.43–1.78); 50–54 h: 0.83 (0.32–2.14); 55–59 h: 0.67 (0.15–2.94); >60 h: 0.48 (0.06–3.68). Cox regression analysis of work hours as a continuous variable: HR=0.93 (0.76–1.13)	Age, sex, marital status, medical specialty, decision authority at work, social support at work, quantitative work demands and previous redemption of AD drug prescription	**Varma 2012**[49]
Mood disorders, GHQ-30	Prevalence of cases by working time (hours/week) ≤ 30 hours: 35.7% (41/115); >30 to 40 hours: 39.0% (39/100); >40 to 50 hours: 37.7% (26/69); >50 hours: 56.8% (46/81); *p*=0.0179 Stepwise multiple logistic regression, >50 h/wk. vs. ≤50 h/wk.: parameter 0.635 (p=0.0293)	Marital status, medical facility, position, and night duty.	**Hayasaka 2007**[44]
Mood disorders, GHQ-12 and SCLDS	No association was found between number of hours worked in a week and depression.	Not adjusted	**Firth Cozens 1987**[42]
Mood disorders, SF-36	Scores when night float vs. standard call at baseline and follow-up: Mental health mean score (SD): baseline = 57.33 (22.63) vs. 65.71 (7.61); follow up = 52.00 (15.49) vs. 60.80 (11.45); p=0.72 Mental health component summary (SD), baseline = 34.84 (14.06) vs. 40.21 (7.61) ; follow up = 30.15 (10.71) vs. 42.40 (6.23); p=0.39 Regression analysis: increased number of hours in hospital correlated with significantly lower SF-36 scores in almost all domains.	Not adjusted	**Zahrai 2011**[50]
Mood disorders, SF-36 and BDI	Improvement of levels before and after duty hours limitations from 90-110 to 78-80 hors/week): Mental health SF-36): no statistically significant improvement Individual questions BDI before and after duty hours limitations, only energy level statistically significant.	Not adjusted	**Stamp 2005**[47]

Health outcome	Main results	Adjustment by confounders	Study ID
Mood disorders, SF-36	Odds ratio (95% CI)for working 47 hours per week or more: Impaired mental health, men: 1.59 (0.95-2.66); women: 1.86 (1.03-3.37)	Age and amount of time in practice	**Sundquist 2000**[48]
General health, SF-36	SF-36: Scores when night float vs. standard call at baseline and follow-up: General health mean score (SD): baseline = 62.11 (17.47) vs. 77.57 (24.25); follow up = 56.43 (24.89) vs. 84.20 (16.50); p=0.41 Physical health component summary (SD): baseline = 46.16 (13.15) vs. 52.01 (13.33); follow up = 39.32 (9.80) vs. 56.15 (2.18); p=0.015 Vitality mean score (SD): baseline = 51.67 (14.58) vs. 51.43 (15.74); follow up = 48.57 (14.92) vs. 51.00 (10.84); p=0.20 Regression analysis demonstrated that the increased number of hours spent in hospital correlated with significantly lower general health, physical function, mental health, role emotional, social function and mental component summary scale scores (all p < 0.05)	Not adjusted	**Zahrai 2011**[50]
General health, SF-36	SF-36, BDI: Improvement of levels before and after duty hours limitations from 90-110 to 78-80 hours/week): Physical health: no statistically significant differences.	Not adjusted	**Stamp 2005**[47]
General health, SF-36	SF-36: Odds ratio (95% CI) adjusted for age and time in practice, for working 47 hours per week or more: Impaired general health: men = 1.66 (1.00-2.77); women = 1.59 (1.00-3.17)	Age and amount of time in practice	**Sundquist 2000**[48]
General health	Self-rated health: Odds ratio (95% CI) of good self-rated health (logistic regression): Norwegian work time pattern**= 1.35 (1.03-1.77); working in Norway = 4.17 (3.02 - 5.73).	Age, sex and country of work	**Rosta 2011**[46]

**Not working more than 9 hours a day and having more than 60 hours a month on-call); GHQ-12= General Health Questionnaire-12 items; SCLDS= Symptom Checklist Depression Scale; BDI = Beck Depression Inventory II.

Percutaneous injuries: One study of high quality and one of intermediate quality, with non-adjusted analyses, showed consistent results in medical trainees. Ayas et al. found an increased risk of percutaneous injuries associated with working more than 20 hours/day compared to non-extended hour periods, except in intensive care units;[40] Fisman et al. found an association between self-reported fatigue and the number of working hours at the time of injury and a 3-fold increased risk of sharp injuries in medical trainees compared to other health care workers (median of working hours/week before the injury: 70 vs. 40;p<.001).[43]

Motor vehicle accidents: Two studies of high and intermediate quality showed that long working weeks were associated with an at least two-fold increased risk. Barger et al. found adjusted increased risks for car crashes (OR 2·3; 95%CI 1·6-3·3) and near miss accidents (OR 5·9; 95%CI 5·4-6·3) associated with working extended shifts.[41] Kirkcaldy et al. showed that non-adjusted incidence rates of traffic accidents on house visits was five times as much when physicians worked for a mean of 58 hours/week compared to 38 hours/week (p<0,05).[45]

Mood disorders: Six studies of intermediate and high quality found contradictory results for mood disorders. Three of them, all of intermediate quality, provided evidence of an association between LWH and mental health problems. The prevalence of GHQ-30 cases was higher for female physicians working above 50 hours/week than for those working 30 or less hours/week (p<0.05),[44] and an increased adjusted risk of impaired mental health was found for men and women general practitioners who worked for 47 hours per week or more.[48] Zahrai et al. found that increased number of hours spent by residents in hospital correlated significantly with lower mental health and mental component summary scale scores (SF-36).[50] Three other studies, one of which of high quality, did not find an association between depression or mood disorders and LWH. Varma et al. did not find higher adjusted risks of depression associated with increased number of hours in senior medical consultants, even when they worked for more than 60 hours per week.[49] Depression or mood disorders measured with GHQ-12, SCLDS or BDI, all validated tools for mental health, were not associated with increased number of working hours [42] nor improved after decreasing the number of hours from 90-110 to 78-80 hours per week in residents.[47]

General health: Four studies of intermediate quality analysed the association between LWH and general health. Zahrai et al. found that increased number of hours spent by

residents in hospital correlated significantly with poorer general health, physical function and vitality using SF-36.[50] A comparative study of physicians in two different countries found that Norwegians showed higher non-adjusted prevalence of self-reported good health compared to physicians in Germany who worked longer hours.[46] An increased adjusted risk of impaired general health was found for men and women general practitioners who worked for more than 47 hours per week.[48] However, following the implementation of reducing the number of hours from an average of 90-110hours/week to 78-80hours/week did not lead to an overall improvement of residents self-reported physical health.[47]

Synthesis of the evidence:

The levels of evidence synthesis obtained from the analysis of the 11 studies included in this review are shown in table 4.

Table 4. Available scientific evidence on the health problems associated with long working hours in physicians: levels of evidence synthesis.

Health outcome	Degree of evidence*	Magnitude of the association**	Studies ID
Percutaneous injuries	++	++	Ayas 2006[40], Fisman 2007[43]
Motor vehicle accidents	++	++	Barger 2005[41], Kirkcaldy 1997[45]
Mood disorders	+	+/-	Varma 2012[49], Hayasaka 2007[44], Firth Cozens 2001[42], Zahrai 2011[50], Stamp 2005[47], Sundquist 2000[48]
General health	+	+/-	Zahrai 2011[50], Stamp 2005[47], Sundquist 2000[48], Rosta 2011[46]

* Strong evidence (+++): consistent results in more than 2 studies of high quality; Moderate evidence (++): consistent results in two studies of high quality, or one high quality study and one intermediate, or between some studies of intermediate quality; Insufficient evidence (+): identification of only one study or inconsistent results across studies; Evidence of no association (-): consistent results of a non-association in two or more studies.

** Magnitude of the association: it refers to the magnitude of the association point estimate (RR = relative risk, rate ratio or odds ratio): high (+++) if RR > 3; intermediate (++) if RR = 1·5-3; low (+) if RR < 1·5; no association (-); unclear (+/-).

There is moderate evidence of an association between LWH and percutaneous and motor vehicle accidents in physicians. This evidence comes from one study of high quality and another of intermediate quality both for percutaneous injuries, and for motor vehicle accidents and the magnitude of this association could be intermediate.

For mood disorders and general health there is a low or insufficient level of evidence of an association with LWH in physicians. This comes from 6 studies on mood disorders and 4 on general health of high and intermediate methodological quality and inconsistent results among them. No conclusion can be drawn about the magnitude of such associations if existed.

DISCUSSION

This review found moderate scientific evidence for a positive association of intermediate magnitude between exposure to LWH and percutaneous injuries and motor vehicle accidents in physicians. Evidence was assessed as low or insufficient for mood disorders and general health.

To our knowledge, this is the first systematic review on the effects of LWH on physicians' health and safety. The extensive searches were restricted to indexed journals. We systematically identified, selected and assessed the methodological quality of studies by means of independent pairs of reviewers. The quality assessment form and the levels of evidence used in this review were based on CONSORT and STROBE statements and on previous systematic reviews.[36, 38, 39] To give some estimate of the potential magnitude of the effect, we added levels based on the point estimates of measures of association provided by the studies. However, such levels had not been previously established.

Strengths and limitations of the studies

Four studies had a longitudinal prospective design, however only two of them used a robust methodology. Two other studies used a crossover design, which reduces the likelihood of individual variability and confounding, and were considered well designed. Four studies, all rated as intermediate methodological quality, were cross sectional in design and therefore it is difficult to draw causal relationships from them.

Although reverse causality cannot be ruled out from cross-sectional designs, it is unlikely that poorer health determines longer working hours than their healthier peers. One study used a quasi-experimental design but lacked a control group. Also, because of different designs used across the included studies, different estimates of the effects were used that may not be directly comparable.

Other limitations include the variability, validity and reliability of working hours, which in most studies was self-reported. Although only those papers specifying the number of working hours were accepted, shift and night work might have worked as confounders too, as it is difficult to acknowledge which proportion of the health effects observed was due exclusively to the LWH component and not due, for instance, to stress, fatigue or sleep deprivation. Moreover, the studies did not discriminate between time spent actively working or asleep while on call, though the EWTD considers all hours on call as working time. Likewise, it is possible that participants were aware that the study was conducted to assess associations between the number of hours worked and health outcomes. Therefore, there might be a volunteer bias in those doctors who work extended hours. To minimise this bias, a case-cross over analysis has been performed in three of the included study [40, 41, 43].

Health outcomes were measured mainly based upon self-report, with the exception of documented motor vehicle crashes,[40] reported percutaneous injuries [42] and register-notified anti-depressive drug prescriptions.[48] No studies with other objective measures of health (e.g., mortality) or mental disease (e.g., hospital data) were identified. However, self-reported or perceived health was assessed by validated and widely used instruments, especially when health status (general, mental or physical health) or ill-health symptoms are evaluated. Nonetheless, this raises the possibility that the observed associations might reflect differences in propensity to report health problems when they occur rather than true differences in the risk of worse health status. We cannot rule this out, but the higher propensity for reporting among those who work for longer hours and poorer work schedule patterns would have to relate specifically to the reporting of ill-health rather than non-ill health. It seems unlikely that major differences in propensity to report would extend to a more concrete outcome.

The majority of the studies of this review took into account confounding variables in their analyses, such as sex and age. However, none of the studies analysed the potential effect of modifying factors, such as psychosocial aspects at work, including attitudes,

motivation, job requirements, demands and content, organisational climate, social relationships at work, work satisfaction, supportive organisation, or the relative number of physicians and other staff available for patient care. Neither were important aspects outside of work, such as life events or lifestyles, taken into account. The study of Japanese female physicians found significant differences by marital status with higher GHQ scores for those who were married.[44] There are significant differences in the number of physicians per capita across different countries and this together with the structure and organisation of health services are potential confounders too that were not addressed in the selected studies across countries.

A relevant outcome from this review would be to establish a threshold number of extended hours above which there is a significantly higher risk. There is a lack of evidence for a dose-response relationship and nor does the evidence give any indication for a threshold number of hours that physicians could work and remain safe and healthy.

Finally, the heterogeneity of the included studies did not allow a more quantitative synthesis, and a level of evidence approach was used instead.

Comparison with other studies

Our findings of an association between LWH and injuries and accidents in physicians are consistent with research in other occupational groups. Dembe et al. carried out a longitudinal survey in the US including more than 10,000 workers from a variety of occupations and settings.[19] A strong dose-response effect (adjusted for age, gender, occupation, industry and region) was found between LWH and injuries above 40h/week and 8h/day; jobs with overtime schedules were associated with a 61% higher injury hazard rate compared to jobs without overtime; working at least 12 hours per day was associated with a 37% increased hazard rate, and working at least 60 hours per week was associated with a 23% increased hazard rate. They concluded that LWH might indirectly precipitate workplace injuries through a causal process by inducing fatigue or stress in workers. In a separate study assessing injury risks to health care personnel, Dembe et al. demonstrated that the risk of injury when working overtime or at least 60 h per week among physicians and nurses was statistically significant.[51] In our review we found that information about the length of shift work varied across studies, some

analysing shifts of up to twenty four hours. Work that included such shifts was more strongly associated with accidents and injuries.[40, 41]

We found insufficient evidence of an association between LWH and mood disorders, other diseases and general health in physicians, despite some scientific evidence suggesting that LWH increase morbidity and mortality in other occupations.[19-26]

Possible explanations for these inconsistencies might be the scarcity of high quality research, and the fact that some of the included papers did not fully address the effect of confounding factors, and the possibility that working as a physician might have a protective effect by itself, for instance through higher motivation and social recognition relative to other occupations.

Stamp et al., in the only before and after study, did not find significant differences for mood disorder or general health before and after implementing requirements to reduce working schedules to 80h/week in residents.[47] It is debatable, however, whether any conclusions can be drawn from the impact of reducing extremely high working hours (90-110hours/week) to very high working hours (78-80hours/week).

Although we searched for and identified literature describing associations between LWH and increased mortality in doctors,[52, 53] none of those studies specified the number of hours worked, and therefore were not included.

Recommendations for future research

Further research, including well designed prospective and retrospective longitudinal studies, is needed in this field to establish on scientific grounds what number of hours are safe, both to patients and physicians. Also studies on long term health effects should be considered.

This review was aimed to study exclusively possible detrimental health effects of LWH (more than 48 hours per week) to physicians. Other systematic reviews should be conducted addressing issues, such as the impact of working hours on doctor's well-being and quality of life, the quality of junior doctors training, the quality and continuity of care and impact on patient safety. Also, from a staffing perspective, the economic and social burden of sick leave and physicians leaving the profession are important issues that were beyond the scope of this review and that should be considered in future

research. Other established criteria [54] that working hours should be 'family friendly', promote gender equality, enhance productivity or facilitate worker choice and influence over working hours, were not addressed. Further research is needed in these other areas.

Conclusions

The findings of this review have confirmed that long working hours are associated with both an increased risk of needlestick injuries and road traffic accidents. The fact that these associations are labelled moderate in this review, should not preclude safety procedures and measures being assessed and reassessed to ensure maximum physician and patient safety. The reduction of needlestick injuries requires improved use of safe sharp devices, compliance with safe working procedures, training and regular audits of working practices, particularly for junior doctors who are at greatest risk of needlestick injuries, possibly due to inadequate training.[55, 56] The risk of road traffic accidents could be reduced by discouraging driving after very long shifts (more than 16 hours) and where there is sleep deprivation.[57] Transport should be provided by the employer in these situations. This is a pragmatic recommendation which recognises that long working hours have not been eliminated as a result of the EWTD.

This review uncovered the lack of literature on the effect of long working hours on the general health of doctors, a topic that has important health ramifications not only for physicians but for patients too. However, the paucity of available studies, which does not allow for a causal or dose-response relationship to be established, raises further questions about the evidence-base of the current 48 hour limit, especially as this is not enforced strictly and junior doctors, for example in the UK, may opt-out.[58] It may be relevant that the EWTD was not based entirely on science but also took into account, as does all EU legislation, the views of "social partners".[3]

It is likely that the mechanism for these increased health and safety risks is fatigue, and our finding of increased risk of accidents to staff, may also be associated with increased risk of clinical errors. Patient safety was not the purpose of this review but long working hours are well recognised to cause decrement in performance both in health care and other professions, where the performance of staff can be safety critical, such as airline pilots, and professional drivers.[59]

In addition the EWTD has been associated with real concerns about the adequacy of the training of doctors given their reduced exposure to patient care, and the lack of

experience of the patient journey, imposed by adherence to the EWTD.[4] This is potentially exacerbated at the same time by the constraints on the overall duration of training imposed by other EU directives.[60] Further reviews of the EWTD, with regard to this unique group of workers, need to take into account other social factors such as the impact on patient care. This is particularly the case when all health care systems are under strain because of burgeoning demands and limited resources.[4]

The findings of this systematic review lead to the suggestion of the following guidelines in the event of non-compliance with the EWTD: 1) long shifts should be avoided to protect both physicians' health and patient safety and no shift should be longer than 16 hours (and then exceptionally);[28] 2) Physicians should be discouraged from driving after long shifts to reduce the risk of motor accidents;[41, 45] 3) rigorous attention should be paid to reducing the risk of sharps injuries;[40, 43] 4) organisational aspects such as workload and job control, as well as the pattern and distribution of working hours, breaks and recovery periods should be carefully taken into account, in order to avoid fatigue and sleep deprivation that could lead to mistakes and accidents;[61, 62] 5) Physicians should work in supportive psychosocial environments, in teams, and with adequate training and supervision as their wellbeing is important also for the health of their patients.[61, 62]

Further research is required to establish any longer term effects such as on mortality and mental health, also the impact of the nature of the work organisation taking into account the psychosocial aspects of the physicians working and non-working lives, and to determine how many hours are safe for physicians to work.

ACKNOWLEDGMENTS

Leena Isotalo, Trial Search Coordinator at the Cochrane Occupational Safety and Health Review Group, designed and conducted the final search of the evidence using Medline and Embase databases.

FUNDING

The Section of Occupational Medicine of the UEMS provided limited support to the work of all authors (MCR, ED, SVP, KS, AS, PR, RH), with the exception of the senior authors (EBM, CS). The expenses of one meeting of the core group (MCR, SVP, ED, EBM, CS) were also funded by the Section. **Role of funding source:** The Section of Occupational Medicine of the UEMS acted in its role of promoting and contributing to the research in Occupational Medicine.

AUTHORS' CONTRIBUTIONS:

Mari Cruz Rodríguez-Jareño, Evangelia Demou, Sergio Vargas-Prada, Alenka Škerjanc, Kaveh A Sanati, Ewan B. Macdonald and Consol Serra conceived and submitted the design, carried out the acquisition of data, analysis and interpretation of data, drafted the article and revised it.

Pedro G. Reis and Ritva Helimäki-Aro contributed to acquisition of data, analysis and interpretation of data, drafted the article and revised it.

All authors gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

CONFLICT OF INTERESTS

None. No financial relationships with any organisations that might have an interest in the submitted work, and no other relationships or activities that could appear to have influenced the submitted work.

DATA SHARING STATEMENT

No additional data are available

ETHICS COMMITTEE APPROVAL was not sought as no patients were involved in this study.

REFERENCES

1 EUR-Lex. Directive 2003/88/EC of the European Parliament and of the Council of 4 November 2003 concerning certain aspects of the organisation of working time. In: Union OJotE, ed. Brussels: Official Journal of the European Union, 2003.

2 EPHA. 2012 EPHA Briefing on the European Working Time Directive (EWTD) http://www.epha.org/IMG/pdf/EPHA_Briefing_EWTD_Oct2012_FINAL.pdf; Access Year 2012.

3 European Commission. 2013 Working Conditions-Working Time Directive. http://ec.europa.eu/social/main.jsp?catId=706&langId=en&intPageId=205; Access Date December 2013.

4 House J. Calling time on doctors' working hours. Lancet 2009;**373**:2011-12.

5 The Lancet. Doctors' training and the European Working Time Directive. Lancet 2010;**375**:2121.

6 Leff D, Aziz O, Darzi A. Trucks, planes, and scalpels - Is there an evidence-based approach to surgeons' working hours? Arch Surg 2007;**142**:817-20.

7 Parthasarathy S. Sleep and the medical profession. Curr Opin Pulm Med 2005;**11**:507-12.

8 Dr. Foster Hospital Guide. 2011 www.drfosterhealth.co.uk; Access Date February 2011.

9 UEMS. 2012 The Union of European Medical Specialists. www.uems.net; Acess Year 2012.

10 Macdonald EB, Ritchie KA, Murray KJ, et al. Requirements for occupational medicine training in Europe: a Delphi study. Occup Environ Med 2000;**57**:98-105.

11 Macdonald E, Baransky B, Wilford J. Occupational Medicine in Europe: Scope and Cometencies. In: Health WECfEa, ed. Health, Environment and Safety in Enterprises Series n. 3. Bilthoven, The Netherlands, 2000.

12 Ballester M, Cordell N, Rodriguez Jareno MC, et al. A European survey of professional bodies representing occupational medicine specialists. Occup Med (Oxf) 2012;**62**:366-70.

13 Cashman C, Slovak A. The occupational medicine agenda: Routes and standards of specialization in occupational medicine in europe. Occup Med (Oxf) 2005;**55**:308-11.

14 De Schryver A, Claesen B, Meheus A, et al. European survey of hepatitis B vaccination policies for healthcare workers. Eur J Public Health 2011;**21**:338-43.

15 Rodriguez-Juareno M, Serra C, Demou E, et al. European Working Time Directive: systematic review for evidence based decision making. SUMMARY REPORT. In:

Medicine USoO, ed. http://www.uems-occupationalmedicine.org/sites/default/files/old_userfiles/File%203%20-%20EWTD%20UEMS%20REPORT%2005%2010%202011.pdf,: UEMS Section of Occupational Medicine, 2011.

16 UEMS. 2010 The Newsletter of European Medical Specialists. http://www.uems.eu/__data/assets/pdf_file/0014/1463/1466.pdf.

17 UEMS. Meetings of the Board & Council of UEMS- Report. In: European Union of Medical Specialists; http://www.uems.eu/__data/assets/pdf_file/0011/1235/UEMS_2011.37_-_report_-_UEMS_Council_6-8.10.2010_Napoli.pdf, ed. Napoli (Italy), 2011.

18 Bannai A, Tamakoshi A. The association between long working hours and health: A systematic review of epidemiological evidence. Scand J Work Environ Health 2013.

19 Dembe AE, Erickson JB, Delbos RG, et al. The impact of overtime and long work hours on occupational injuries and illnesses: new evidence from the United States. Occup Environ Med 2005;**62**:588-97.

20 Johnson JV, Lipscomb J. Long working hours, occupational health and the changing nature of work organization. Am J Ind Med 2006;**49**:921-9.

21 Kivimaeki M, Batty GD, Hamer M, et al. Using Additional Information on Working Hours to Predict Coronary Heart Disease. Ann Intern Med 2011;**154**.

22 Liu Y, Tanaka H, Fukuoka Heart Study G. Overtime work, insufficient sleep, and risk of non-fatal acute myocardial infarction in Japanese men. Occup Environ Med 2002;**59**:447-51.

23 Sokejima S, Kagamimori S. Working hours as a risk factor for acute myocardial infarction in Japan: case-control study. BMJ 1998;**317**:775-80.

24 Spurgeon A, Harrington JM, Cooper CL. Health and safety problems associated with long working hours: A review of the current position. Occup Environ Med 1997;**54**:367-75.

25 Virtanen M, Ferrie JE, Singh-Manoux A, et al. Long working hours and symptoms of anxiety and depression: a 5-year follow-up of the Whitehall II study. Psychol Med 2011;**41**:2485-94.

26 Virtanen M, Stansfeld SA, Fuhrer R, et al. Overtime work as a predictor of major depressive episode: a 5-year follow-up of the Whitehall II study. PloS One 2012;**7**:e30719.

27 Gopal R, Glasheen JJ, Miyoshi TJ, et al.. Burnout and internal medicine resident work-hour restrictions. Arch Intern Med 2005;**165**:2595-600.

28 Reed DA, Fletcher KE, Arora VM. Systematic Review: Association of Shift Length, Protected Sleep Time, and Night Float With Patient Care, Residents' Health, and Education. Ann Intern Med 2010;**153**:829-42.

29 Sargent MC, Sotile W, Sotile MO, et al.. Quality of Life During Orthopedic Training and Academic Practice Part 1: Orthopedic Surgery Residents and Faculty. J Bone Joint Surg Am 2009;**91A**:2395-405.

30 Fletcher KE, Underwood W, Davis SQ, et al. Effects of work hour reduction on residents' lives - A systematic review. JAMA 2005;**294**:1088-100.

31 Black C, Frost D. *Health at work - an independent review on sickness absence*. London: The Stationary Office, 2011.

32 Waddell G, Burton K. *Is work good for your health and wellbeing?* London: The Stationary Office, 2006.

33 Butterworth P, Leach LS, Strazdins L, et al. The psychosocial quality of work determines whether employment has benefits for mental health: results from a longitudinal national household panel survey. Occup Environ Med 2011;**68**:806-12.

34 Lockley SW, Landrigan CP, Barger LK, et al. When policy meets physiology - The challenge of reducing resident work hours. Clin Orthop 2006:116-27.

35 ICD-10. 2012 International Classification of Diseases. http://www.who.int/classifications/icd/en/; Access Year 2012.

36 van Uffelen JGZ, Wong J, Chau JY, et al. Occupational Sitting and Health Risks A Systematic Review. Am J Prev Med 2010;**39**:379-88.

37 Kirsling RA, Kochar MS, Chan CH. An evaluation of mood states among 1st-year residents. Psychol Rep 1989;**65**:355-66.

38 Bernard B. A critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity and low back. In: NIOSH, ed. Publication No. 9741. Cincinnati: National Institute for Occupational Safety and Health, 1997.

39 Steenstra IA, Verbeek JH, Heymans MW, et al. Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature. Occup Environ Med 2005;**62**:851-60.

40 Ayas NT, Barger LK, Cade BE, et al. Extended work duration and the risk of self-reported percutaneous injuries in interns. JAMA 2006;**296**:1055-62.

41 Barger LK, Cade BE, Ayas NT, et al. Extended work shifts and the risk of motor vehicle crashes among interns. N Engl J Med 2005;**352**:125-34.

42 Firth-Cozens J. Emotional Distress in Junior House Officers. BMJ 1987;**295**:533-36.

43 Fisman DN, Harris AD, Rubin M, et al. Fatigue increases the risk of injury from sharp devices in medical trainees: Results from a case-crossover study. Infect Control HospEpidemiology 2007;**28**:10-17.

44 Hayasaka Y, Nakamura K, Yamamoto M, et al. Work environment and mental health status assessed by the general health questionnaire in female Japanese doctors. Ind Health 2007;**45**:781-86.

45 Kirkcaldy BD, Trimpop R, Cooper CL. Working hours, job stress, work satisfaction, and accident rates among medical practitioners and allied personnel. Int J Stress Manag 1997;**4**:79-87.

46 Rosta J, Aasland OG. Work Hours and Self rated Health of Hospital Doctors in Norway and Germany. A comparative study on national samples. BMC Health Serv Res 2011;**11**.

47 Stamp T, Termuhlen P, Miller S, et al. Before and after resident work hour limitations: an objective assessment of the well-being of surgical residents. Curr Surg 2005;**62**:117-21.

48 Sundquist J, Johansson SE. High demand, low control, and impaired general health: working conditions in a sample of Swedish general practitioners. Scand J Pub Health 2000;**28**:123-31.

49 Varma A, Marott J, Stoltenberg C, et al. With long hours of work, might depression then lurk? A nationwide prospective follow-up study among Danish senior medical consultants. Scand J Work Environ Health 2012;**38**:418-26.

50 Zahrai A, Chahal J, Stojimirovic D, et al. Quality of life and educational benefit among orthopedic surgery residents: a prospective, multicentre comparison of the night float and the standard call systems. Can J Surg 2011;**54**:25-32.

51 Dembe AE, Delbos R, Erickson JB. Estimates of injury risks for healthcare personnel working night shifts and long hours. Quality & Safety in Health Care 2009;**18**:336-40.

52 Lindfors PM, Nurmi KE, Meretoja OA, et al. On-call stress among Finnish anaesthetists. Anaesthesia 2006;**61**:859-64.

53 Šelb J, Albreht T. Mortality rates of medical doctors in Slovenia in 1985 to 1999. Zdravniški vestnik 2000:I47-I48.

54 Lee S, McCann D, Messenger J. Working Time Around the World: Trends in working hours, laws and policies in a global comparative perspective. Geneva, Switzerland: International Labour Office, 2007.

55 Elder A, Paterson C. Sharps injuries in UK health care: a review of injury rates, viral transmission and potential efficacy of safety devices. Occup Med-Oxford 2006;**56**:566-74.

56 Naghavi SHR, Sanati KA. Needlestick injuries: Does left-handedness matter? Am J Infect Control 2009;**37**:341-41.

57 Connor J, Norton R, Ameratunga S, et al. Driver sleepiness and risk of serious injury to car occupants: population based case control study. BMJ 2002;**324**:1125-28A.

58 BMA. 2013 European Working Time Directive, http://bma.org.uk/practical-support-at-work/ewtd,. Access Date December 2013.

59 Wilson AM, Weston G. Application of Airline Pilots Hours to Junior Doctors. BMJ 1989;**299**:779-81.

60 The European Specialist Medical Qualifications Order 1995. 1995 Statutory Instruments, 1995 No. 3208, MEDICAL PROFESSION, The European Specialist Medical Qualifications Order 1995, http://legislation.data.gov.uk/uksi/1995/3208/made/data.htm?wrap=true,. Access Date December 2013.

61 Boorman S. NHS Health and Well-being. The Boorman Review. Leeds, 2009.

62 Williams S, Michie S, Pattani S. Improving the health of the NHS workforce. Report of the partnership on the health of the NHS Workforce. London: The Nuffield Trust, 1998.

Figure legend

Figure 1. Results of the search strategy, using search engines on PubMed and EMBASE (December 2011) and screening of references lists of identified full papers, study selection and quality assessment.

European Working Time Directive and doctors' health. A Systematic review of the available epidemiological evidence.

Running title: Long working hours and doctors' health

***Corresponding author**

Dr. Evangelia Demou
Healthy Working Lives Group
Institute of Health and Wellbeing
University of Glasgow
Glasgow, G12 8RZ
Tel: +44 141 330 3559
E-mail: evangelia.demou@glasgow.ac.uk

Maria Cruz Rodriguez-Jareño[1,2], Evangelia Demou[3,*], Sergio Vargas-Prada[4], Kaveh A Sanati[3,5], Alenka Škerjanc[6], Pedro G. Reis[7], Ritva Helimäki-Aro[8], Ewan B. Macdonald[3], Consol Serra[1,4,9,10], on behalf of the UEMS Section of Occupational Medicine[11].

1. Catalan and Spanish Societies of Occupational Medicine, Barcelona, Spain.
2. Department of Medical Sciences. School of Medicine. Universitat de Girona. Girona. Spain
3. Institute of Health and Wellbeing, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom
4. CiSAL - Centre for Research in Occupational Health, Universitat Pompeu Fabra, Barcelona, Spain.
5. Occupational Health Department, Epsom & St Helier University Hospitals NHS Trust, Surrey, United Kingdom
6. Clinical Institute of Occupational, Traffic and Sports Medicine, University Medical Centre, Ljubljana, Slovenia.
7. College of Occupational Medicine, Portuguese Medical Association, Lisbon, Portugal.
8. Helsinki City Occupational Health Centre, Helsinki, Finland.
9. Occupational Health Service, Parc de Salut MAR, Barcelona, Spain.
10. CIBER of Epidemiology and Public Health. Spain.
11. UEMS Section of Occupational Medicine. http://www.uems-occupationalmedicine.org

Word Count: 4,~~667~~794

ABSTRACT

Objective: To summarise the available scientific evidence on the health effects of exposure to working beyond the limit number of hours established by the European Working Time Directive (EWTD) on physicians.

Design: A systematic literature search was conducted in PubMed and Embase. Study selection, quality appraisal and data extraction were carried out by independent pairs of researchers using pre-established criteria.

Setting: Physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.

Participants: The total number of participants was 14,338.

Primary and secondary outcome measures: Health effects classified under the International Classification of Diseases (ICD-10).

Results: Over 3,000 citations and 110 full articles were reviewed. From these, eleven studies of high or intermediate quality carried out in North America, Europe and Japan met the inclusion criteria. Six studies included medical residents, junior doctors or house officers and the five others included medical specialists or consultants, medical, dental and general practitioners and hospital physicians. Evidence of an association was found between percutaneous injuries and road traffic accidents with extended hours/day or very long working weeks (LWH). The evidence was insufficient for mood disorders and general health. No studies on other health outcomes were identified.

Conclusions: LWH could increase the risk of percutaneous injuries and road traffic accidents, and possibly other incidents at work through the same pathway. While associations are clear, the existing evidence does not allow for an established causal or "dose-response" relationship between LWH and incidents at work, or for a threshold number of extended hours above which there is a significantly higher risk and the hours physicians could work and remain safe and healthy. Policy makers should consider safety issues when working on relaxing EWTD for doctors.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- The first systematic review, to our knowledge, on the effects of Long Working Hours (LWH) on physicians' health and safety.
- A systematic literature search conducted in PubMed and Embase with over 3,000 citations and 110 full articles reviewed.
- Eleven studies of high or intermediate quality carried out in North America, Europe and Japan, involving physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.
- The findings of this review confirmed that long working hours are associated with both an increased risk of needlestick injuries and road traffic accidents. Evidence was assessed as low or insufficient for mood disorders and general health.
- This review uncovered the lack of literature on the effect of long working hours on the general health of doctors, a topic that has important health ramifications not only for physicians but for patients too.

KEY WORDS: Working time, "Physicians"[MeSH], "Morbidity"[MeSH], "Occupational Injuries"[MeSH], Systematic review.

INTRODUCTION

The European Working Time Directive (EWTD),[1] applicable to all occupations across the EU, requires a maximum working week of 48 hours and establishes rest periods.[2, 3] Since it came into force in healthcare in 2009, it has been associated with concerns about the provision of health services including continuity of care, lower staffing levels, introduction of shift working, a reduction in training time and the adequate supervision of junior doctors.[4, 5] The large inter-country variations in physicians working hours as well as the mandatory controls over work hours across occupations raises further questions as to the basis of selecting and setting these limits and restrictions.[6] There is evidence of variable compliance to the EWTD within health care across Europe.[4, 5] The medical profession is also increasingly feminised with potential implications for women of childbearing age, which may coincide with the time in the profession where long working hours (LWH) are more likely to happen (i.e. during residency). Constraint on public-sector finances and increasing health care demands, have stimulated the debate on physicians' working hours,[7] with some evidence showing that the quality of health care is positively correlated with the number of physicians available to deliver that care.[8]

The Union of European Medical Specialists (UEMS) was created in 1958 to represent medical specialists in the EU, promote a high standard of training and practice at European level and facilitate the free movement of physicians across European countries.[9] The Occupational Medicine section was created in 1997 and is involved in issues related to specialist training, professional practice and research;[10] elaborating and publishing reference documents,[11] conducting and contributing to surveys,[12-14] and working in partnership with other key European groups.

The trigger for this project was a request by the Council of the UEMS to its Occupational Medicine Section for a position statement on the EWTD and possible health consequences to physicians of a potential removal of this professional group from the current EWTD 48 hours per week limit~~. An evidence based approach was taken by the Section and a systematic review of the literature was undertaken.~~[15-17] An evidence based approach was taken by the Section and a systematic review of the literature was undertaken.[15]

LWH have been associated with several adverse effects on workers' health, such as hypertension, cardiovascular disease, stress, depression, musculoskeletal disorders,

chronic infections, diabetes, general health complaints and all-cause mortality in a variety of occupational groups.~~[15-23]~~[18-26] In physicians, there is some evidence suggesting that LWH can result in impacts such as stress, depression, burnout, injuries, fatigue and sleep deprivation,~~[24-26]~~[27-29] and overwork has been considered one of the most stressful features of physicians' work.~~[24]~~[27] Some evidence exists of the beneficial effect of reducing the number of working hours on the quality of life of medical residents ~~[27]~~[30] and on burnout.[~~24~~27]

Field Code Changed
Field Code Changed

However, the relationship between work and health is complex. There is a substantial body of evidence showing that worklessness is associated with poorer health, whereas work is generally good for health and well-being,[31, 32] provided that the work environment is reasonably acceptable and supportive.[33]

Occupational and non-occupational exposures may play an important role such as the pattern and distribution of working hours, breaks, and recovery periods. Also, psychosocial and organisational factors, such as workload, job control, managers and peer support, training opportunities, individual characteristics and attitudes may be important.[6, 34]

Finally, physicians are a highly qualified, devoted and motivated professional group. Despite other existing reviews of a more general occupational approach, a review focused on physicians is warranted.

The aim of this project was to systematically review whether long-working hours (LWH) -i.e. defined as more than the 48 hour per week limit imposed by the EWTD- are associated with health effects –i.e. classified under the International Classification of Diseases (ICD-10) ~~[32]~~[35]- in physicians, and to examine what these associations are. This review does not include the potential impact of LWH on patient care or physician training.

METHODS

The systematic review sought to respond to the following research question: are LWH associated with health effects in physicians? For the purpose of this review, exposure to LWH was defined as working for more than 48 hours per week.~~[1]~~[1] Health outcomes included any disease as defined in the ICD-10 ~~[32] and work related injuries.~~[35] and work related injuries. If self-assessed tools on health status were used, only those validated were considered. Burnout, stress and outcomes such as individual symptoms, signs or biological markers (blood pressure, ECG, etc.) were not considered. Physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.

Study identification

Electronic searches were carried out using PubMed and EMBASE as search engines (December 2011). Our search strategy was similar in both databases and consisted of a combination of the following keywords and Mesh terms: night shift, morning shift, evening shift, afternoon shift, shift work, rotating shift, shift combination, shift duration or length, shift system, clockwise rotation, shift roster, extended shifts, night work, evening work, work schedule, work hours, starting time, early start, irregular working hours, direction of rotation, overwork, extended hours, shift rota, workload, work schedule tolerance, sleep deprivation, sleep disorders, chronobiology disorders, circadian rhythm, psychomotor performance, circadian disruption, vigilance, alertness, wakefulness, drowsiness, fatigue, insomnia, hypersomnolence, dyssomnia, eveningness, morningness, neurocognitive performance, concentration difficulties, arousal, health, morbidity, mortality, disease, illness, stress, strain, distress, accident, injur*, death, suicid*, education, medical, physician, medical staff, hospital, doctor, surgeon, house officers, medical school, surgery, surgical.

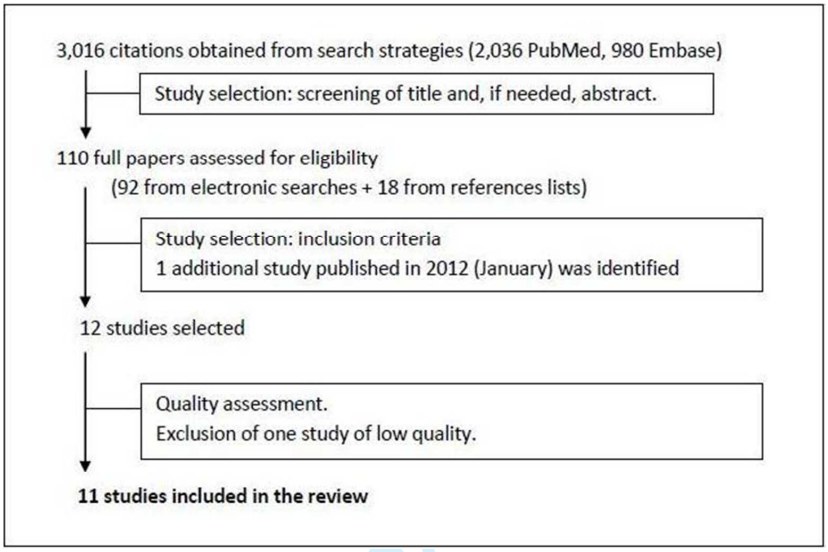
In addition, the references lists of articles selected for inclusion were carefully reviewed to identify additional studies.

Study selection

Studies were included if they measured the association between the exposure to LWH and health effects in physicians, used an observational epidemiological (i.e. cohort, case-control, cross-sectional) or experimental design~~, and were published in English, French, German, Italian, Slovenian or Spanish~~. Articles relating to on-call, night or shift work, but with unknown exposure or exposed to less than 48 working hours per week were excluded. They were also excluded if the working groups were other than physicians or addressed only other exposures or outcomes (i.e. patient safety, fatigue, sleep deprivation, social or family disruption). No limitation was set for languages.

A total of 2,036 citations were obtained from PubMed and 980 non-duplicated citations using Embase, yielding a total of 3,016 citations that were all screened by independent pairs of reviewers. All titles and, where necessary, abstracts were reviewed. Ninety two potentially suitable publications were identified from the electronic searches, and 18 further studies were identified from the references lists, yielding a total of 110 studies ~~for all of which~~published in English, French, German, Italian, Slovenian and Spanish. For all of them the full text was obtained and reviewed by independent pairs of reviewers. Disagreements within pairs were resolved by discussion and, where necessary, by a third reviewer who made the final decision. All authors participated as reviewers for screening the citations and full papers using well defined and pre-established criteria. Ninety nine articles were excluded at this stage. One further study was identified that was published after the search period. Finally, 12 papers were considered for quality assessment (Figure 1).

Figure 1. Results of the search strategy, using search engines on PubMed and EMBASE (December 2011) and screening of references lists of identified full papers, study selection and quality assessment.



The methodological quality of the 12 studies was assessed by independent pairs of reviewers. A standardized 16 point scale based on CONSORT and STROBE statements (Table 1) and adapted from a previous systematic review ~~[33] was used.~~ [36] was used. It includes 16 items grouped into 6 areas: 1) objectives, 2) study design, 3) target population and sample, 4) variables, 5) data sources, collection and measurement, and 6) statistical methods. Each item was rated as 1 (the requirement was met), 0.5 (the requirement was partially met) or 0 (the requirement was not or unclearly met). Disagreements within pairs of reviewers were resolved by consensus or, where necessary, by a third reviewer. For each of the 12 studies, a final score based on the sum of all items was assigned and the percentage was calculated based on a maximum score of 16. Study quality was rated as low, moderate or high if it scored less than 60%, between 60% and 79.9%, and 80% or more of the maximum score, respectively.

~~One article of low quality was excluded [34] and a total of 11 studies of intermediate or high methodological quality were included for the purpose of this review (figure 1).~~

One article of low quality was excluded [37] and a total of 11 studies of intermediate or high methodological quality were included for the purpose of this review (figure 1).

Data extraction

Selected information was extracted from each paper, including publication year, country of origin, study design, setting, study population, sample size, response rate, measure of working hours, definition used for LWH, health outcomes and their measurement, and

main results on point risk estimates or frequencies of health outcomes, and their corresponding 95% confidence interval (95%CI), and whether the analyses were adjusted for potential confounders.

Evidence synthesis

To summarise the results on the relationship between LWH and health effects, levels of evidence synthesis was performed. This was based on the methodological quality, study design and the consistency of the study outcomes. The following criteria were based on two previous relevant systematic reviews.[~~35~~38, ~~36~~39]

Field Code Changed

Field Code Changed

Strong evidence: consistent results in more than two studies of high quality.

Moderate evidence: consistent results in one high quality study and one intermediate, or between some studies of intermediate quality.

Insufficient evidence: identification of only one study or inconsistent results across studies;

Evidence of no association: consistent results of a non-association in two or more studies.

A measure of the possible magnitude of the association was attempted using the following criteria based on the association point estimate (RR = relative risk, rate ratio or odds ratio) when available or otherwise the comparison between two frequencies was used:

- high: RR > 3·00
- intermediate: RR = 1·50-3·00
- low: RR = 1·01-1·49
- no association: RR=1·00
- unclear

RESULTS

We identified 11 studies related to LWH and health effects in physicians that were eligible for inclusion. The outcome of the methodological quality assessment is given in table 1.

Methodological quality was appraised as high for 3 of the 11 studies, with scores ranging from 84 to 94%. The other 8 studies were considered as intermediate quality as

their scores ranged from 63 to 78%. The majority of all included studies received positive scores on items describing the study objectives, design and the study population (items from areas 1, 2 and 3), as well as the description of exposure and outcome variables (item 4). However, although data sources and collection were well described, reliability and validity of exposure to LWH and health outcome measures had low scores (items 5b and 5c). The statistical methods were in general appropriate, however the confounding variables were not measured in several studies and, as can be seen, the studies did not describe how they addressed missing data (items 6a and 6b).

Study characteristics:

The characteristics of the included studies are described in table 2.

Table 1. Methodological quality appraisal of the included studies

Study ID			Ayas 2006 [37][40]	Barger 2005 [38][41]	Firth-Cozens 1987 [39][42]	Fisman 2007 [40][43]	Hayasaka 2007 [41][44]	Kirkcaldy 1997 [42][45]	Rosta 2011 [43][46]	Stamp 2005 [44][47]	Sundquist 2000 [45][48]	Varma 2012 [46][49]	Zahrai 2011 [47][50]
1	Objectives	Are the objectives or hypotheses of the research described in the paper?	1	1	1	1	1	1	1	1	1	1	1
2	Study design	Is the study design presented?	1	1	1	1	1	1	1	1	1	1	1
3a	Target population	Do the authors describe the target population they wanted to research?	1	1	1	1	1	1	1	1	1	1	1
3b	Sample	Was a random sample of the target population taken? AND was the response rate 60% or more?	0.5	0.5	1	0.5	0.5	0.5	0.5	1	1	1	0.5
3c	Sample	Is participant selection described?	1	1	1	1	0.5	1	0	1	1	1	1
3d	Sample	Is participant recruitment described, or referred to?	1	1	1	1	1	1	1	1	1	1	0
3e	Sample	Are the inclusion and/or exclusion criteria stated?	1	1	1	1	0.5	0.5	1	1	1	1	0.5
3f	Sample	Is the study sample described? (minimum description: sample size, gender, age and occupation)	1	1	0	1	1	1	0	0	0	1	1
3g	Sample	Is the number of participants at each stage of the study reported?	1	1	1	0.5	1	0.5	1	1	1	1	1
4	Variables	Are the measures of long working hours and the health outcome described?	1	1	0.5	1	1	1	1	1	1	1	1
5a	Data sources, collection	Do authors describe the source of their data (e.g., official registry, health survey) AND how the data were collected?	1	1	1	1	1	1	1	1	1	1	1
5b	Measurement	Was reliability of the measure(s) of long working hours mentioned or referred to?	0	0	0	1	0	0	0	1	0	1	0
5c	Measurement	Was the validity of the measure(s) of long working hours mentioned or referred to?	1	1	0	0	0	0	0	0	0	0	0
5d	Measurement	Were health outcomes assessed by objective measures or validated self-reporting instruments?	0	1	1	1	1	0	1	1	1	1	1

6a	Statistical methods	Were appropriate statistical methods used and described, including those for addressing confounders?	0.5	1	0.5	1	0.5	0.5	1	0.5	1	1	0.5
6b	Statistical methods	Were the numbers/% of participants with missing data for long working hours and the health outcome indicated AND If more than 20% of data in the primary analyses were missing, were methods used to address missing data?	0	0	0	0.5	0	0	0	0	0	1	0.5
Total score** (%)			12.0 (75)	13.5 (84)	11.0 (69)	13.5 (84)	11.0 (69)	10.0 (63)	10.5 (66)	12.5 (78)	12.0 (75)	15.0 (94)	11.0 (69)
Quality rate***			interm	high	interm	high	interm	interm	interm	interm	interm	high	interm

** Maximum score = 16; *** Quality rate (%): low=<60; intermediate (interm)=60-79; high=80-100.

Table 2. Characteristics of the included studies.

Study ID	Country Of Study Population	Design*	Setting	Participants and sample size (response rate)	Working hours	Health outcomes	Quality score (%)
~~Ayas 2006~~ ~~[37]~~Ayas 2006 [40]	United States	~~Cross-over~~Cohort, prospective (1 year) / Case-crossover	Hospital	Interns in postgraduate residency programs Sample size: 2,737 (80%)	Mean (SD) hrs worked/month: 249.8 (75.3~~:self~~). Self-reported, monthly survey. Strong correlation (Pearson *r*=0.76; *p* 0.001) with hours worked 244 (69.3) from work diaries of randomly selected subset of 192 interns. Extended periods ~~(~~>(20 ~~hrs/day~~or more consecutive hours) vs. non extended periods (~~first~~ 12 hrs ~~at work~~or less consecutive hours).	Self-reported percutaneous injuries	12.0 (75)
~~Barger 2005~~ ~~[38]~~Barger 2005 [41]	United States	Cohort, prospective (1 year)	Hospital	Residents first postgraduate year (interns). Sample size: 2,~~554 (75~~737 (80%)	~~Average 70.7±26.0~~Mean (SD) hrs worked/week~~.~~ 70.7 (26.0). Extended shift (≥ 24hrs) vs. non extended shift (<24 hours). Self-reported, validated.	Self-reported and documented motor vehicle crashes.	13,5 (84)
~~Firth Cozens 1987~~ ~~[39]~~Firth Cozens 1987 [42]	United Kingdom	Cohort, prospective (1 year)	Hospital	Junior house officers Sample size: 170 (72%)	Mean number of hours per week = 90.6 hours (include on call). Self-reported.	GHQ-12 (case: score ≥2) and SCLDS, self-reported	11.0 (69)
~~Fisman 2007~~ ~~[40]~~Fisman 2007 [43]	United States and Canada	~~Cross-over~~Case-Crossover	Hospital	Medical trainees vs. other HCW. Sample size: 109 vs. 241 (46%)	Median number of hours per week: medical trainees = 70; other HCW = 40 (p<0.001). Self-reported, high reliability.	Reported percutaneous injuries to employee health care service.	13.5 (84)
~~Hayasaka 2007~~ ~~[41]~~Hayasaka 2007 [44]	Japan	Cross-sectional	Hospital, clinics, other	Female physicians Sample size: 367 (63%)	Comparison of increasing number of hrs/week from ≤ 30 to >50, self-reported.	GHQ-30 (case: score ≥8), self-reported	11.0 (69)
~~Kirkcaldy 1997~~ ~~[42]~~Kirkcaldy 1997 [45]	Germany	Cross-sectional	Not specified	Medical and dental practitioners. Sample size: 2,500 (not specified)	Long hours: mean 58.36 hrs/week SD 9.16 Short hours: mean 38.17 hrs/week SD 7.72 Self-reported.	Self-reported traffic accidents	10.0 (63)
~~Rosta 2011~~ ~~[43]~~Rosta 2011 [46]	Germany and Norway	Cross-sectional	Hospital	Hospital physicians. ~~Sample size~~Participants: 1,917 (58%) in Germany and 1,072 (65%) in Norway answered the questionnaire. From survey responders, 1,260 (65.7%) Germans and 562 (52.4) Norwegians ~~(58% and 65%,)~~were included in the analysis after applying further inclusion criteria.	German vs. Norwegian physicians (%): Hours per day >9 = 58.8 vs. 26.7 60 hours on-call per month = 63.4 vs. 18.3 Self-reported.	Validated questionnaire on self-rated health	10.5 (66)

Formatted: Font: (Intl) Arial Unicode MS
Formatted: Font: (Intl) Arial
Formatted: Font: Font color: Black, (Intl) Arial
Formatted: Font: Font color: Black, (Intl) Arial
Formatted: Font: Font color: Black, (Intl) Arial
Formatted: Font: Font color: Black, (Intl) Arial

~~Stamp 2005~~ [44]Stamp 2005 [47]	United States	Before-and-after	Hospital	Residents of general surgery Sample size: 28 (97%)	Changes of work patterns, after reduction from 90-110 to 78-80 work hrs/week.	SF-36, mental health; BDI, depression; self-reported	12.5 (78)
~~Sundquist 2000~~ [45]Sundquist 2000 [48]	Sweden	Cross-sectional	Primary care	General practitioners Sample size: 1,004 (72%)	Overtime defined as working at least 47 hrs/week. Self-reported.	Swedish SF-36, impaired mental health, self-reported	12.0 (75)
~~Varma 2012~~ [46]Varma 2012 [49]	Denmark	Cohort, prospective (20 months)	Hospital	Senior medical consultants Sample size: 2,790 (62%)	Long work hours (>40hrs/week). Self-reported.	Depression: redemption of anti-depressive drug prescriptions	15.0 (94)
~~Zahrai 2011~~ [47]Zahrai 2011 [50]	Canada	Cohort, prospective (6 months)	Hospital	Orthopaedic surgery residents Sample size: 16 (not specified)	Night float (n=9): 77.8% did >80h/week at baseline; 71.4% at follow-up). Standard call (n=7): 57.1% did >80h week at baseline; 80% at follow-up.	SF-36, mental health score, self-reported	11.0 (69)

* Follow-up period in brackets for prospective cohort studies; GHQ-12= General Health Questionnaire-12 items; SCLDS= Symptom Checklist Depression Scale; BDI = Beck Depression Inventory II.

Five had been carried out in North America (4 in the United States ~~[37, 38, 40, 44]~~[40, 41, 43, 47] and 1 in Canada [50]), five in European countries (Denmark,[49] Germany,[45] Norway and Germany [43]; Sweden,[48] and the United Kingdom [42]) and one study in Japan.[44] Four used a prospective cohort design (follow-up ranging from 6 to 20 months), 2 were case-crossover, 4 cross-sectional, and 1 was a before and after study. Overall, the total number of participants was 14,338 and included medical residents, junior doctors or house officers only working in hospitals, medical specialists or consultants, medical and dental practitioners, general practitioners, hospital physicians and one study included female physicians from a community service. Most studies (n=9) compared different working hour patterns within the same group of physicians, one compared physicians from two different countries,[46] and one included other health care workers as the comparison group.[43]

LWH was defined as more than 48hours/week in most studies, and some did not provide a definition. Two studies considered LWH below 48/week.[46, 48] All the included studies investigated health effects associated with working more than 48h/week, with number of hours ranging between <30 to 110hours/week. One paper studied the effects of an increasing number of hours from 30 to more than 50h/week.[44] Others studied more than 40h/week,[49] 47h/week,[48] above 58h/week [46] (two studies) and 6 studies referred to more than 70h/week,~~[37 40, 44, 47]~~[40-43, 47, 50] one of which compared 78-80h/week to 90-110 weekly hours.[47] Information on working hours was self-reported in nine studies, with reasonably good validity only in two studies ~~[37, 38]~~[40, 41] and reliability in one.[43] In the other two studies it was obtained from established work schedules.~~[44, 47]~~[47, 50]

Identified health outcomes that fulfilled the inclusion criteria were percutaneous injuries, motor vehicle accidents, mood disorders and general health. Information was self-reported in all studies using well known validated questionnaires, except three studies that used documented information on motor vehicle crashes,[40] reported percutaneous injuries [42] and antidepressant prescription data as a surrogate of depression.[49]

Summary of findings:

Table 3 shows the findings from each of the 11 studies.

Table 3. Results on the association between long working hours and identified health effects in physicians.

Health outcome	Main results	Adjustment by confounders	Study ID
Percutaneous injuries	Odds ratio (95% CI) for injuries during extended (>20 hours/day) vs. non extended periods (first 12 hours at work): All percutaneous injuries: 1.61 (1.46-1.78); Injuries reported to OH: 1.83 (1.48-2.28); Injuries in the ICU: 1.87 (0.69-5.04); Injuries in the operating room or labour and delivery: 1.77 (1.49-2.09); Injuries in the ICU, non-ICU, or ED: 2.17 (1.56-3.00)	Not adjustedTime of day and circadian influences	Ayas 2006[37]Ayas 2006[40]
Percutaneous injuries	Medical trainees vs. other health care workers: total median working hours: 70 vs. 40 hours (p<.001); median previous working hours at the time of injury: 6.5 vs. 5 hours (p< .001). Odds ratio (95% CI) for self-reported fatigue>: work > 40h in the past week: 3.550 (2.06-5.92); work during more than 5d in past week: 4.220 (4.46-7.15); >12h at work before injury: 8.58 (3.7-19.86). Incidence rate ratio (95%CI) for association between fatigue and injuries and fatigue: all workers: 1.40 (1.03-1.90); medical trainees: 2.94 (1.71-5.07); other HCWs: 0.97 (0.66-1.42) (p =0.001 for heterogeneity).	Age and sex	Fisman 2007[40]Fisman 2007[43]
Motor vehicle accidents	Relative riskOdds ratios (95% CI), after extended shift (≥ 24hrs) vs. non extended shift (<24 hours): Crashes = 2.3 (1.6-3.3); Near miss accidents = 5.9 (5.4–6.3)	Age and sex	Barger 2005[38]Barger 2005[41]
Motor vehicle accidents	Incidence rates of driving accidents on house visits: Long hours (mean 58.36 hrs/week) = 0.10 (0.31) vs. short hours (mean 38.17 hrs/week) = 0.02 (0.18) vs.). p<0.05	Not adjusted	Kirkcaldy 1997[42]Kirkcaldy 1997[45]
Mood disorders, depression	Hazard ratio (95% CI) hours/week intervals and redemption of anti-depressive drug prescription(reference group: 37-40 h/week): 25–36 h: 0.83 (0.24–2.82); 41–44 h: 0.95 (0.50–1.77); 45–49 h: 0.88 (0.43–1.78); 50–54 h: 0.83 (0.32–2.14); 55–59 h: 0.67 (0.15–2.94); >60 h: 0.48 (0.06–3.68). Cox regression analysis of work hours as a continuous variable: HR=0.93 (0.76–1.13)	Age, sex, marital status, medical specialty, decision authority at work, social support at work, quantitative work demands and previous redemption of AD drug prescription	Varma 2012[46]Varma 2012[49]
Mood disorders, GHQ-30	Prevalence of cases by working time (hours/week) ≤ 30 hours: 35.7% (41/115); >30 to 40 hours: 39.0% (39/100); >40 to 50 hours: 37.7% (26/69); >50 hours: 56.8% (46/81); *p*=0.0179 Stepwise multiple logistic regression, >50 h/wk. vs. ≤50 h/wk.: parameter 0.635 (p=0.0293)	Marital status, medical facility, position, and night duty.	Hayasaka 2007[41]Hayasaka 2007[44]
Mood disorders, GHQ-12 and SCLDS	No association was found between number of hours worked in a week and depression.	Not adjusted	Firth Cozens 1987[39]Firth Cozens 1987[42]
Mood disorders, SF-36	Scores when night float vs. standard call at baseline and follow-up: Mental health mean score (SD): baseline = 57.33 (22.63) vs. 65.71 (7.61); follow up = 52.00 (15.49) vs. 60.80 (11.45); p=0.72 Mental health component summary (SD), baseline = 34.84 (14.06) vs. 40.21 (7.61) ; follow up = 30.15 (10.71) vs. 42.40 (6.23); p=0.39 Regression analysis: increased number of hours in hospital correlated with significantly lower SF-36 scores in almost all domains.	Not adjusted	Zahrai 2011[47]Zahrai 2011[50]

Health outcome	Main results	Adjustment by confounders	Study ID
Mood disorders, SF-36 and BDI	Improvement of levels before and after duty hours limitations from 90-110 to 78-80 hors/week): Mental health SF-36): no statistically significant improvement Individual questions BDI before and after duty hours limitations, only energy level statistically significant.	Not adjusted	~~Stamp 2005~~[44]~~Stamp 2005~~[47]
Mood disorders, SF-36	Odds ratio (95% CI)for working 47 hours per week or more: Impaired mental health, men: 1.59 (0.95-2.66); women: 1.86 (1.03-3.37)	Age and amount of time in practice	~~Sundquist 2000~~[45]~~Sundquist 2000~~[48]
General health, SF-36	SF-36: Scores when night float vs. standard call at baseline and follow-up: General health mean score (SD): baseline = 62.11 (17.47) vs. 77.57 (24.25); follow up = 56.43 (24.89) vs. 84.20 (16.50); p=0.41 Physical health component summary (SD): baseline = 46.16 (13.15) vs. 52.01 (13.33); follow up = 39.32 (9.80) vs. 56.15 (2.18); p=0.015 Vitality mean score (SD): baseline = 51.67 (14.58) vs. 51.43 (15.74); follow up = 48.57 (14.92) vs. 51.00 (10.84); p=0.20 Regression analysis demonstrated that the increased number of hours spent in hospital correlated with significantly lower general health, physical function, mental health, role emotional, social function and mental component summary scale scores (all p < 0.05)	Not adjusted	~~Zahrai 2011~~[47]~~Zahrai 2011~~[50]
General health, SF-36	SF-36, BDI: Improvement of levels before and after duty hours limitations from 90-110 to 78-80 hours/week): Physical health: no statistically significant differences.	Not adjusted	~~Stamp 2005~~[44]~~Stamp 2005~~[47]
General health, SF-36	SF-36: Odds ratio (95% CI) adjusted for age and time in practice, for working 47 hours per week or more: Impaired general health: men = 1.66 (1.00-2.77); women = 1.59 (1.00-3.17)	Age and amount of time in practice	~~Sundquist 2000~~[45]~~Sundquist 2000~~[48]
General health~~, SF-36~~	Self-rated health: Odds ratio (95% CI) of good self-rated health (logistic regression): Norwegian work time pattern**= 1.35 (1.03-1.77); working in Norway = 4.17 (3.02 - 5.73).	Age, sex and country of work	~~Rosta 2011~~[43]~~Rosta 2011~~[46]

**Not working more than 9 hours a day and having more than 60 hours a month on-call); GHQ-12= General Health Questionnaire-12 items; SCLDS= Symptom Checklist Depression Scale; BDI = Beck Depression Inventory II.

Percutaneous injuries: One study of high quality and one of intermediate quality, with non-adjusted analyses, showed consistent results in medical trainees. Ayas et al. found an increased risk of percutaneous injuries associated with working more than 20 hours/day compared to non-extended hour periods, except in intensive care units;~~[37] Fisman et al.~~[40] Fisman et al. found an association between self-reported fatigue and the number of working hours at the time of injury and a 3-fold increased risk of sharp injuries in medical trainees compared to other health care workers (median of working hours/week before the injury: 70 vs. 40;p<.001).[~~40~~43]

Motor vehicle accidents: Two studies of high and intermediate quality showed that long working weeks were associated with an at least two-fold increased risk. Barger et al. found adjusted increased risks for car crashes (OR 2·3; 95%CI 1·6-3·3) and near miss accidents (OR 5·9; 95%CI 5·4-6·3) associated with working extended shifts.~~[38] Kirkcaldy et al.~~[41] Kirkcaldy et al. showed that non-adjusted incidence rates of traffic accidents on house visits was five times as much when physicians worked for a mean of 58 hours/week compared to 38 hours/week (p<0,05).[~~42~~45]

Mood disorders: Six studies of intermediate and high quality found contradictory results for mood disorders. Three of them, all of intermediate quality, provided evidence of an association between LWH and mental health problems. The prevalence of GHQ-30 cases was higher for female physicians working above 50 hours/week than for those working 30 or less hours/week (p<0.05),~~[41]~~[44] and an increased adjusted risk of impaired mental health was found for men and women general practitioners who worked for 47 hours per week or more.~~[45] Zahrai et al.~~[48] Zahrai et al. found that increased number of hours spent by residents in hospital correlated significantly with lower mental health and mental component summary scale scores (SF-36).~~[47]~~[50] Three other studies, one of which of high quality, did not find an association between depression or mood disorders and LWH. Varma et al. did not find higher adjusted risks of depression associated with increased number of hours in senior medical consultants, even when they worked for more than 60 hours per week.~~[46]~~[49] Depression or mood disorders measured with GHQ-12, SCLDS or BDI, all validated tools for mental health, were not associated with increased number of working hours ~~[39]~~[42] nor improved after decreasing the number of hours from 90-110 to 78-80 hours per week in residents.[~~44~~47]

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

General health: Four studies of intermediate quality analysed the association between LWH and general health. Zahrai et al. found that increased number of hours spent by residents in hospital correlated significantly with poorer general health, physical function and vitality using SF-36.~~[47]~~[50] A comparative study of physicians in two different countries found that Norwegians showed higher non-adjusted prevalence of self-reported good health compared to physicians in Germany who worked longer hours.~~[43]~~[46] An increased adjusted risk of impaired general health was found for men and women general practitioners who worked for more than 47 hours per week.~~[45]~~[48] However, following the implementation of reducing the number of hours from an average of 90-110hours/week to 78-80hours/week did not lead to an overall improvement of residents self-reported physical health.[~~44~~47]

Field Code Changed

Field Code Changed

Synthesis of the evidence:

The levels of evidence synthesis obtained from the analysis of the 11 studies included in this review are shown in table 4.

Table 4. Available scientific evidence on the health problems associated with long working hours in physicians: levels of evidence synthesis.

Health outcome	Degree of evidence*	Magnitude of the association**	Studies ID
Percutaneous injuries	++	++	~~Ayas 2006[37], Fisman 2007[40]~~Ayas 2006[40], Fisman 2007[43]
~~MMotor~~Motor vehicle accidents	++	++	~~Barger 2005[38], Kirkcaldy 1997[42]~~Barger 2005[41], Kirkcaldy 1997[45]
Mood disorders	+	+/-	~~Varma 2012[46], Hayasaka 2007[41], Firth Cozens 2001[39], Zahrai 2011[47], Stamp 2005[44], Sundquist 2000[45]~~Varma 2012[49], Hayasaka 2007[44], Firth Cozens 2001[42], Zahrai 2011[50], Stamp 2005[47], Sundquist 2000[48]
General health	+	+/-	~~Zahrai 2011[47], Stamp 2005[44], Sundquist 2000[45], Rosta 2011[43]~~Zahrai 2011[50], Stamp 2005[47], Sundquist 2000[48], Rosta 2011[46]

Formatted: Indent: Hanging: 0.1"

* Strong evidence (+++): consistent results in more than 2 studies of high quality; Moderate evidence (++): consistent results in two studies of high quality, or one high quality study and one intermediate, or between some studies of intermediate quality; Insufficient evidence (+): identification of only one

study or inconsistent results across studies; Evidence of no association (-): consistent results of a non-association in two or more studies.

** Magnitude of the association: it refers to the magnitude of the association point estimate (RR = relative risk, rate ratio or odds ratio): high (+++) if RR > 3; intermediate (++) if RR = 1·5-3; low (+) if RR < 1·5; no association (-); unclear (+/-).

There is moderate evidence of an association between LWH and percutaneous and motor vehicle accidents in physicians. This evidence comes from one study of high quality and another of intermediate quality both for percutaneous injuries, and for motor vehicle accidents and the magnitude of this association could be intermediate.

For mood disorders and general health there is a low or insufficient level of evidence of an association with LWH in physicians. This comes from 6 studies on mood disorders and 4 on general health of high and intermediate methodological quality and inconsistent results among them. No conclusion can be drawn about the magnitude of such associations if existed.

DISCUSSION

This review found moderate scientific evidence for a positive association of intermediate magnitude between exposure to LWH and percutaneous injuries and motor vehicle accidents in physicians. Evidence was assessed as low or insufficient for mood disorders and general health.

To our knowledge, this is the first systematic review on the effects of LWH on physicians' health and safety. The extensive searches were restricted to indexed journals. We systematically identified, selected and assessed the methodological quality of studies by means of independent pairs of reviewers. The quality assessment form and the levels of evidence used in this review were based on CONSORT and STROBE statements and on previous systematic reviews.~~[33, 35, 36]~~[36, 38, 39] To give some estimate of the potential magnitude of the effect, we added levels based on the point estimates of measures of association provided by the studies. However, such levels had not been previously established.

Strengths and limitations of the studies

Four studies had a longitudinal prospective design, however only two of them used a robust methodology. Two other studies used a crossover design, which reduces the likelihood of individual variability and confounding, and were considered well designed. Four studies, all rated as intermediate methodological quality, were cross sectional in design and therefore it is difficult to draw causal relationships from them. Although reverse causality cannot be ruled out from cross-sectional designs, it is unlikely that poorer health determines longer working hours than their healthier peers. One study used a quasi-experimental design but lacked a control group. Also, because of different designs used across the included studies, different estimates of the effects were used that may not be directly comparable.

Other limitations include the variability, validity and reliability of working hours, which in most studies was self-reported. Although only those papers specifying the number of working hours were accepted, shift and night work might have worked as confounders too, as it is difficult to acknowledge which proportion of the health effects observed was due exclusively to the LWH component and not due, for instance, to stress, fatigue or sleep deprivation. Moreover, the studies did not discriminate between time spent actively working or asleep while on call, though the EWTD considers all hours on call as working time. Likewise, it is possible that participants were aware that the study was conducted to assess associations between the number of hours worked and health outcomes. Therefore, there might be a volunteer bias in those doctors who work extended hours. To minimise this bias, a case-cross over analysis has been performed in three of the included study [40, 41, 43].

Formatted: Font: (Intl) Arial Unicode MS

Health outcomes were measured mainly based upon self-report, with the exception of documented motor vehicle crashes,~~[37]~~[40] reported percutaneous injuries ~~[39]~~[42] and register-notified anti-depressive drug prescriptions.~~[45] No studies with other objective measures of health (e.g.,~~[48] No studies with other objective measures of health (e.g., mortality) or mental disease (e.g., hospital data) were identified. However, self-reported or perceived health was assessed by validated and widely used instruments, especially when health status (general, mental or physical health) or ill-health symptoms are evaluated. Nonetheless, this raises the possibility that the observed associations might reflect differences in propensity to report health problems when they occur rather than true differences in the risk of worse health status. We cannot rule this out, but the higher propensity for reporting among those who work for longer hours and poorer work

schedule patterns would have to relate specifically to the reporting of ill-health rather than non-ill health. It seems unlikely that major differences in propensity to report would extend to a more concrete outcome.

The majority of the studies of this review took into account confounding variables in their analyses, such as sex and age. However, none of the studies analysed the potential effect of modifying factors, such as psychosocial aspects at work, including attitudes, motivation, job requirements, demands and content, organisational climate, social relationships at work, work satisfaction, supportive organisation, or the relative number of physicians and other staff available for patient care. Neither were important aspects outside of work, such as life events or lifestyles, taken into account. The study of Japanese female physicians found significant differences by marital status with higher GHQ scores for those who were married.[~~41~~44] There are significant differences in the number of physicians per capita across different countries and this together with the structure and organisation of health services are potential confounders too that were not addressed in the selected studies across countries.

Field Code Changed

Field Code Changed

A relevant outcome from this review would be to establish a threshold number of extended hours above which there is a significantly higher risk. There is a lack of evidence for a dose-response relationship and nor does the evidence give any indication for a threshold number of hours that physicians could work and remain safe and healthy.

Finally, the heterogeneity of the included studies did not allow a more quantitative synthesis, and a level of evidence approach was used instead.

Comparison with other studies

Our findings of an association between LWH and injuries and accidents in physicians are consistent with research in other occupational groups. Dembe et al. carried out a longitudinal survey in the US including more than 10,000 workers from a variety of occupations and settings.~~[16]~~[19] A strong dose-response effect (adjusted for age, gender, occupation, industry and region) was found between LWH and injuries above 40h/week and 8h/day; jobs with overtime schedules were associated with a 61% higher injury hazard rate compared to jobs without overtime; working at least 12 hours per day was associated with a 37% increased hazard rate, and working at least 60 hours per week was associated with a 23% increased hazard rate. They concluded that LWH

might indirectly precipitate workplace injuries through a causal process by inducing fatigue or stress in workers. In a separate study assessing injury risks to health care personnel, Dembe et al. demonstrated that the risk of injury when working overtime or at least 60 h per week among physicians and nurses was statistically significant.[51] In our review we found that information about the length of shift work varied across studies, some analysing shifts of up to twenty four hours. Work that included such shifts was more strongly associated with accidents and injuries.[~~37~~40, ~~38~~41]

Field Code Changed

We found insufficient evidence of an association between LWH and mood disorders, other diseases and general health in physicians, despite some scientific evidence suggesting that LWH increase morbidity and mortality in other occupations.[~~16 23~~19-26]

Field Code Changed

Possible explanations for these inconsistencies might be the scarcity of high quality research, and the fact that some of the included papers did not fully address the effect of confounding factors, and the possibility that working as a physician might have a protective effect by itself, for instance through higher motivation and social recognition relative to other occupations.

Stamp et al., in the only before and after study, did not find significant differences for mood disorder or general health before and after implementing requirements to reduce working schedules to 80h/week in residents.~~[44]~~[47] It is debatable, however, whether any conclusions can be drawn from the impact of reducing extremely high working hours (90-110hours/week) to very high working hours (78-80hours/week).

~~Although we searched for and identified literature describing associations between LWH and increased mortality in doctors,[48, 49] none of those studies specified the number of hours worked, and therefore were not included.~~

Although we searched for and identified literature describing associations between LWH and increased mortality in doctors,[52, 53] none of those studies specified the number of hours worked, and therefore were not included.

Recommendations for future research

Further research, including well designed prospective and retrospective longitudinal studies, is needed in this field to establish on scientific grounds what number of hours

are safe, both to patients and physicians. Also studies on long term health effects should be considered.

This review was aimed to study exclusively possible detrimental health effects of LWH (more than 48 hours per week) to physicians. Other systematic reviews should be conducted addressing issues, such as the impact of working hours on doctor's well-being and quality of life, the quality of junior doctors training, the quality and continuity of care and impact on patient safety. Also, from a staffing perspective, the economic and social burden of sick leave and physicians leaving the profession are important issues that were beyond the scope of this review and that should be considered in future research. ~~Other established criteria [50]~~Other established criteria [54] that working hours should be 'family friendly', promote gender equality, enhance productivity or facilitate worker choice and influence over working hours, were not addressed. Further research is needed in these other areas.

Conclusions

The findings of this review have confirmed that long working hours are associated with both an increased risk of needlestick injuries and road traffic accidents. The fact that these associations are labelled moderate in this review, should not preclude safety procedures and measures being assessed and reassessed to ensure maximum physician and patient safety. The reduction of needlestick injuries requires improved use of safe sharp devices, compliance with safe working procedures, training and regular audits of working practices, particularly for junior doctors who are at greatest risk of needlestick injuries, possibly due to inadequate training.~~[51, 52]~~[55, 56] The risk of road traffic accidents could be reduced by discouraging driving after very long shifts (more than 16 hours) and where there is sleep deprivation.~~[53]~~[57] Transport should be provided by the employer in these situations. This is a pragmatic recommendation which recognises that long working hours have not been eliminated as a result of the EWTD.

This review uncovered the lack of literature on the effect of long working hours on the general health of doctors, a topic that has important health ramifications not only for physicians but for patients too. However, the paucity of available studies, which does not allow for a causal or dose-response relationship to be established, raises further questions about the evidence-base of the current 48 hour limit, especially as this is not

enforced strictly and junior doctors, for example in the UK, may opt-out.[54][58] It may be relevant that the EWTD was not based entirely on science but also took into account, as does all EU legislation, the views of "social partners".[3][3]

It is likely that the mechanism for these increased health and safety risks is fatigue, and our finding of increased risk of accidents to staff, may also be associated with increased risk of clinical errors. Patient safety was not the purpose of this review but long working hours are well recognised to cause decrement in performance both in health care and other professions, where the performance of staff can be safety critical, such as airline pilots, and professional drivers.[5559]

Field Code Changed

Field Code Changed

In addition the EWTD has been associated with real concerns about the adequacy of the training of doctors given their reduced exposure to patient care, and the lack of experience of the patient journey, imposed by adherence to the EWTD.[4][4] This is potentially exacerbated at the same time by the constraints on the overall duration of training imposed by other EU directives.[56][60] Further reviews of the EWTD, with regard to this unique group of workers, need to take into account other social factors such as the impact on patient care. This is particularly the case when all health care systems are under strain because of burgeoning demands and limited resources.[4][4]

The findings of this systematic review lead to the suggestion of the following guidelines in the event of non-compliance with the EWTD: 1) long shifts should be avoided to protect both physicians' health and patient safety and no shift should be longer than 16 hours (and then exceptionally);[28] 2) Physicians should be discouraged from driving after long shifts to reduce the risk of motor accidents;[41, 45] 3) rigorous attention should be paid to reducing the risk of sharps injuries;[40, 43] 4) organisational aspects such as workload and job control, as well as the pattern and distribution of working hours, breaks and recovery periods should be carefully taken into account, in order to avoid fatigue and sleep deprivation that could lead to mistakes and accidents;[61, 62] 5) Physicians should work in supportive psychosocial environments, in teams, and with adequate training and supervision as their wellbeing is important also for the health of their patients.[61, 62]

Further research is required to establish any longer term effects such as on mortality and mental health, also the impact of the nature of the work organisation taking into account

the psychosocial aspects of the physicians working and non-working lives, and to determine how many hours are safe for physicians to work.

ACKNOWLEDGMENTS

Leena Isotalo, Trial Search Coordinator at the Cochrane Occupational Safety and Health Review Group, designed and conducted the final search of the evidence using Medline and Embase databases.

CONFLICT OF INTERESTS

None. No financial relationships with any organisations that might have an interest in the submitted work, and no other relationships or activities that could appear to have influenced the submitted work.

FUNDING

The Section of Occupational Medicine of the UEMS provided limited support to the work of all authors (MCR, ED, SVP, KS, AS, PR, RH), with the exception of the senior authors (EBM, CS). The expenses of one meeting of the core group (MCR, ED, SVP, EBM, CS) were also funded by the Section. **Role of funding source:** The Section of Occupational Medicine of the UEMS acted in its role of promoting and contributing to the research in Occupational Medicine.

AUTHORS' CONTRIBUTIONS:

Mari Cruz Rodríguez-Jareño, Evangelia Demou, Sergio Vargas-Prada, Alenka Škerjanc, Kaveh A Sanati, Ewan B. Macdonald and Consol Serra conceived and submitted the design, carried out the acquisition of data, analysis and interpretation of data, drafted the article and revised it.

Pedro G. Reis and Ritva Helimäki-Aro contributed to acquisition of data, analysis and interpretation of data, drafted the article and revised it.

All authors gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

ETHICS COMMITTEE APPROVAL was not sought as no patients were involved in this study.

REFERENCES

1 EUR-Lex. Directive 2003/88/EC of the European Parliament and of the Council of 4 November 2003 concerning certain aspects of the organisation of working time. In: Union OJotE, ed. Brussels: Official Journal of the European Union, 2003.

2 EPHA. 2012 EPHA Briefing on the European Working Time Directive (EWTD) http://www.epha.org/IMG/pdf/EPHA_Briefing_EWTD_Oct2012_FINAL.pdf; Access Year 2012.

3 European Commission. 2013 Working Conditions-Working Time Directive. http://ec.europa.eu/social/main.jsp?catId=706&langId=en&intPageId=205; Access Date December 2013.

4 House J. Calling time on doctors' working hours. Lancet 2009;**373**:2011-12.

5 The Lancet. Doctors' training and the European Working Time Directive. Lancet 2010;**375**:2121.

6 Leff D, Aziz O, Darzi A. Trucks, planes, and scalpels - Is there an evidence-based approach to surgeons' working hours? Arch Surg 2007;**142**:817-20.

7 Parthasarathy S. Sleep and the medical profession. Curr Opin Pulm Med 2005;**11**:507-12.

8 Dr. Foster Hospital Guide. 2011 www.drfosterhealth.co.uk; Access Date February 2011.

9 UEMS. 2012 The Union of European Medical Specialists. www.uems.net; Acess Year 2012.

10 Macdonald EB, Ritchie KA, Murray KJ, Gilmour WH. Requirements for occupational medicine training in Europe: a Delphi study. Occup Environ Med 2000;**57**:98-105.

11 Macdonald E, Baransky B, Wilford J. Occupational Medicine in Europe: Scope and Cometencies. In: Health WECfEa, ed. Health, Environment and Safety in Enterprises Series n. 3. Bilthoven, The Netherlands, 2000.

12 Ballester M, Cordell N, Rodriguez Jareno MC, Serra C, Med USO. A European survey of professional bodies representing occupational medicine specialists. Occup Med (Oxf) 2012;**62**:366-70.

13 Cashman C, Slovak A. The occupational medicine agenda: Routes and standards of specialization in occupational medicine in europe. Occup Med (Oxf) 2005;**55**:308-11.

14 De Schryver A, Claesen B, Meheus A, van Sprundel M, Francois G. European survey of hepatitis B vaccination policies for healthcare workers. Eur J Public Health 2011;**21**:338-43.

~~15~~15 Rodriguez-Juareno M, Serra C, Demou E, et al. European Working Time Directive: systematic review for evidence based decision making. SUMMARY REPORT. In: Medicine USoO, ed. http://www.uems-occupationalmedicine.org/sites/default/files/old_userfiles/File%203%20-%20EWTD%20UEMS%20REPORT%2005%2010%202011.pdf,: UEMS Section of Occupational Medicine, 2011.

16 UEMS. 2010 The Newsletter of European Medical Specialists. http://www.uems.eu/__data/assets/pdf_file/0014/1463/1466.pdf.

17 UEMS. Meetings of the Board & Council of UEMS- Report. In: European Union of Medical Specialists; http://www.uems.eu/__data/assets/pdf_file/0011/1235/UEMS_2011.37_-_report_-_UEMS_Council_6-8.10.2010_Napoli.pdf, ed. Napoli (Italy), 2011.

18 Bannai A, Tamakoshi A. The association between long working hours and health: A systematic review of epidemiological evidence. Scand J Work Environ Health 2013.

~~16~~19 Dembe AE, Erickson JB, Delbos RG, Banks SM. The impact of overtime and long work hours on occupational injuries and illnesses: new evidence from the United States. Occup Environ Med 2005;**62**:588-97.

~~17~~20 Johnson JV, Lipscomb J. Long working hours, occupational health and the changing nature of work organization. Am J Ind Med 2006;**49**:921-9.

~~18~~21 Kivimaeki M, Batty GD, Hamer M, et al. Using Additional Information on Working Hours to Predict Coronary Heart Disease. Ann Intern Med 2011;**154**.

~~19~~22 Liu Y, Tanaka H, Fukuoka Heart Study G. Overtime work, insufficient sleep, and risk of non-fatal acute myocardial infarction in Japanese men. Occup Environ Med 2002;**59**:447-51.

~~20~~23 Sokejima S, Kagamimori S. Working hours as a risk factor for acute myocardial infarction in Japan: case-control study. BMJ 1998;**317**:775-80.

~~21~~24 Spurgeon A, Harrington JM, Cooper CL. Health and safety problems associated with long working hours: A review of the current position. Occup Environ Med 1997;**54**:367-75.

~~22~~25 Virtanen M, Ferrie JE, Singh-Manoux A, et al. Long working hours and symptoms of anxiety and depression: a 5-year follow-up of the Whitehall II study. Psychol Med 2011;**41**:2485-94.

~~23~~26 Virtanen M, Stansfeld SA, Fuhrer R, Ferrie JE, Kivimaki M. Overtime work as a predictor of major depressive episode: a 5-year follow-up of the Whitehall II study. PloS One 2012;**7**:e30719.

2427 Gopal R, Glasheen JJ, Miyoshi TJ, Prochazka AV. Burnout and internal medicine resident work-hour restrictions. Arch Intern Med 2005;**165**:2595-600.

2528 Reed DA, Fletcher KE, Arora VM. Systematic Review: Association of Shift Length, Protected Sleep Time, and Night Float With Patient Care, Residents' Health, and Education. Ann Intern Med 2010;**153**:829-42.

2629 Sargent MC, Sotile W, Sotile MO, Rubash H, Barrack RL. Quality of Life During Orthopedic Training and Academic Practice Part 1: Orthopedic Surgery Residents and Faculty. J Bone Joint Surg Am 2009;**91A**:2395-405.

2730 Fletcher KE, Underwood W, Davis SQ, Mangrulkar RS, McMahon LF, Saint S. Effects of work hour reduction on residents' lives - A systematic review. JAMA 2005;**294**:1088-100.

2831 Black C, Frost D. *Health at work - an independent review on sickness absence*. London: The Stationary Office, 2011.

2932 Waddell G, Burton K. *Is work good for your health and wellbeing?* London: The Stationary Office, 2006.

3033 Butterworth P, Leach LS, Strazdins L, Olesen SC, Rodgers B, Broom DH. The psychosocial quality of work determines whether employment has benefits for mental health: results from a longitudinal national household panel survey. Occup Environ Med 2011;**68**:806-12.

3134 Lockley SW, Landrigan CP, Barger LK, Czeisler CA. When policy meets physiology - The challenge of reducing resident work hours. Clin Orthop 2006:116-27.

3235 ICD-10. 2012 International Classification of Diseases. http://www.who.int/classifications/icd/en/; Access Year 2012.

3336 van Uffelen JGZ, Wong J, Chau JY, et al. Occupational Sitting and Health Risks A Systematic Review. Am J Prev Med 2010;**39**:379-88.

3437 Kirsling RA, Kochar MS, Chan CH. An evaluation of mood states among 1st-year residents. Psychol Rep 1989;**65**:355-66.

3538 Bernard B. A critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity and low back. In: NIOSH, ed. Publication No. 9741. Cincinnati: National Institute for Occupational Safety and Health, 1997.

3639 Steenstra IA, Verbeek JH, Heymans MW, Bongers PM. Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature. Occup Environ Med 2005;**62**:851-60.

3740 Ayas NT, Barger LK, Cade BE, et al. Extended work duration and the risk of self-reported percutaneous injuries in interns. JAMA 2006;**296**:1055-62.

~~38~~41 Barger LK, Cade BE, Ayas NT, et al. Extended work shifts and the risk of motor vehicle crashes among interns. N Engl J Med 2005;**352**:125-34.

~~39~~42 Firth-Cozens J. Emotional Distress in Junior House Officers. BMJ 1987;**295**:533-36.

~~40~~43 Fisman DN, Harris AD, Rubin M, Sorock GS, Mittleman MA. Fatigue increases the risk of injury from sharp devices in medical trainees: Results from a case-crossover study. Infect Control HospEpidemiology 2007;**28**:10-17.

~~41~~44 Hayasaka Y, Nakamura K, Yamamoto M, Sasaki S. Work environment and mental health status assessed by the general health questionnaire in female Japanese doctors. Ind Health 2007;**45**:781-86.

~~42~~45 Kirkcaldy BD, Trimpop R, Cooper CL. Working hours, job stress, work satisfaction, and accident rates among medical practitioners and allied personnel. Int J Stress Manag 1997;**4**:79-87.

~~43~~46 Rosta J, Aasland OG. Work Hours and Self rated Health of Hospital Doctors in Norway and Germany. A comparative study on national samples. BMC Health Serv Res 2011;**11**.

~~44~~47 Stamp T, Termuhlen P, Miller S, et al. Before and after resident work hour limitations: an objective assessment of the well-being of surgical residents. Curr Surg 2005;**62**:117-21.

~~45~~48 Sundquist J, Johansson SE. High demand, low control, and impaired general health: working conditions in a sample of Swedish general practitioners. Scand J Pub Health 2000;**28**:123-31.

~~46~~49 Varma A, Marott J, Stoltenberg C, Wieclaw J, Kolstad H, Bonde J. With long hours of work, might depression then lurk? A nationwide prospective follow-up study among Danish senior medical consultants. Scand J Work Environ Health 2012;**38**:418-26.

~~47~~50 Zahrai A, Chahal J, Stojimirovic D, Schemitsch EH, Yee A, Kraemer W. Quality of life and educational benefit among orthopedic surgery residents: a prospective, multicentre comparison of the night float and the standard call systems. Can J Surg 2011;**54**:25-32.

~~48~~51 Dembe AE, Delbos R, Erickson JB. Estimates of injury risks for healthcare personnel working night shifts and long hours. Quality & Safety in Health Care 2009;**18**:336-40.

52 Lindfors PM, Nurmi KE, Meretoja OA, et al. On-call stress among Finnish anaesthetists. Anaesthesia 2006;**61**:859-64.

~~49~~53 Šelb J, Albreht T. Mortality rates of medical doctors in Slovenia in 1985 to 1999. Zdravniški vestnik 2000:I47-I48.

~~50~~54 Lee S, McCann D, Messenger J. Working Time Around the World: Trends in working hours, laws and policies in a global comparative perspective. Geneva, Switzerland: International Labour Office, 2007.

~~51~~55 Elder A, Paterson C. Sharps injuries in UK health care: a review of injury rates, viral transmission and potential efficacy of safety devices. Occup Med-Oxford 2006;**56**:566-74.

~~52~~56 Naghavi SHR, Sanati KA. Needlestick injuries: Does left-handedness matter? Am J Infect Control 2009;**37**:341-41.

~~53~~57 Connor J, Norton R, Ameratunga S, et al. Driver sleepiness and risk of serious injury to car occupants: population based case control study. BMJ 2002;**324**:1125-28A.

~~54~~58 BMA. 2013 European Working Time Directive, http://bma.org.uk/practical-support-at-work/ewtd,. Access Date December 2013.

~~55~~59 Wilson AM, Weston G. Application of Airline Pilots Hours to Junior Doctors. BMJ 1989;**299**:779-81.

~~56~~60 The European Specialist Medical Qualifications Order 1995. 1995 Statutory Instruments, 1995 No. 3208, MEDICAL PROFESSION, The European Specialist Medical Qualifications Order 1995, http://legislation.data.gov.uk/uksi/1995/3208/made/data.htm?wrap=true,. Access Date December 2013.

~~57~~61 Boorman S. NHS Health and Well-being. The Boorman Review. Leeds, 2009.

~~58~~62 Williams S, Michie S, Pattani S. Improving the health of the NHS workforce. Report of the partnership on the health of the NHS Workforce. London: The Nuffield Trust, 1998.

Formatted: EndNote Bibliography, Indent: Left: 0", Hanging: 0.19", Space Before: 0 pt, After: 0 pt, Line spacing: single

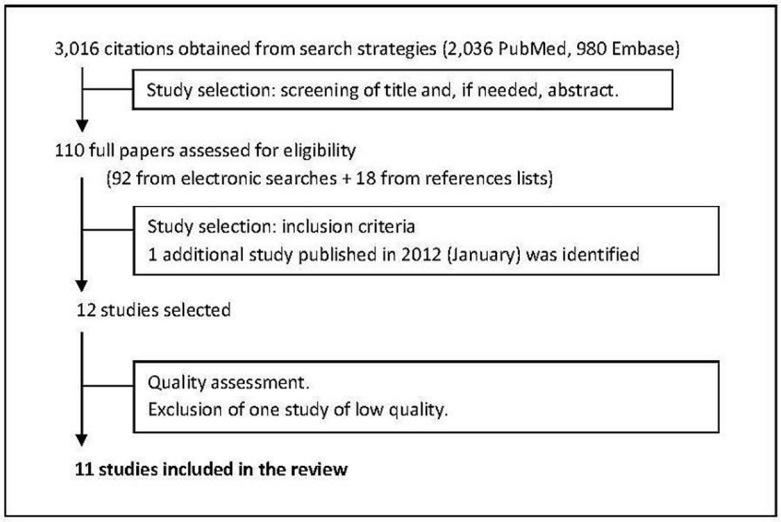


Figure 1.Results of the search strategy, using search engines on PubMed and EMBASE (December 2011) and screening of references lists of identified full papers, study selection and quality assessment.
90x59mm (300 x 300 DPI)

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6-7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	6
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	8-9
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	8-9
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	8-9
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8-9
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	9
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	8-9

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	7
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	10-11
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	12-13
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	18
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	9-19
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	19
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	19-21
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	21-23
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	24-25

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.

BMJ Open

European Working Time Directive and doctors' health. A Systematic review of the available epidemiological evidence.

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-004916.R2
Article Type:	Research
Date Submitted by the Author:	12-Jun-2014
Complete List of Authors:	Rodriguez-Jareno, Maria-Cruz; Universitat de Girona, 2. Department of Medical Sciences Demou, Evangelia; University of Glasgow, Institute of Health and Wellbeing Vargas, Sergio; Universitat Pompeu Fabra, Centre for Research in Occupational Health Sanati, Kaveh; University of Glasgow, Institute of Health and Wellbeing Skerjanc, Alenka; University Medical Centre, 6. Clinical Institute of Occupational, Traffic and Sports Medicine Reis, Pedro; 7. College of Occupational Medicine, Portuguese Medical Association Helimaki-Aro, Ritva; 8. Helsinki City Occupational Health Centre, Macdonald, Ewan; University of Glasgow, Institute of Health and Wellbeing Serra, Consol; Universitat Pompeu Fabra, Centre for Research in Occupational Health
Primary Subject Heading:	Health policy
Secondary Subject Heading:	Medical management, Research methods, Health services research, Health policy, Evidence based practice
Keywords:	EPIDEMIOLOGY, GENERAL MEDICINE (see Internal Medicine), HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

European Working Time Directive and doctors' health. A Systematic review of the available epidemiological evidence.

Running title: Long working hours and doctors' health

***Corresponding author**

Dr. Evangelia Demou
Healthy Working Lives Group
Institute of Health and Wellbeing
University of Glasgow
Glasgow, G12 8RZ
Tel: +44 141 330 3559
E-mail: evangelia.demou@glasgow.ac.uk

Maria Cruz Rodriguez-Jareño[1,2], Evangelia Demou[3,*], Sergio Vargas-Prada[4], Kaveh A Sanati[3,5], Alenka Škerjanc[6], Pedro G. Reis[7], Ritva Helimäki-Aro[8], Ewan B. Macdonald[3], Consol Serra[1,4,9,10], on behalf of the UEMS Section of Occupational Medicine[11].

1. Catalan and Spanish Societies of Occupational Medicine, Barcelona, Spain.
2. Department of Medical Sciences. School of Medicine. Universitat de Girona. Girona. Spain
3. Institute of Health and Wellbeing, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom
4. CiSAL - Centre for Research in Occupational Health, Universitat Pompeu Fabra, Barcelona, Spain.
5. Occupational Health Department, Epsom & St Helier University Hospitals NHS Trust, Surrey, United Kingdom
6. Clinical Institute of Occupational, Traffic and Sports Medicine, University Medical Centre, Ljubljana, Slovenia.
7. College of Occupational Medicine, Portuguese Medical Association, Lisbon, Portugal.
8. Helsinki City Occupational Health Centre, Helsinki, Finland.
9. Occupational Health Service, Parc de Salut MAR, Barcelona, Spain.
10. CIBER of Epidemiology and Public Health. Spain.
11. UEMS Section of Occupational Medicine. http://www.uems-occupationalmedicine.org

Word Count: 4,794

ABSTRACT

Objective: To summarise the available scientific evidence on the health effects of exposure to working beyond the limit number of hours established by the European Working Time Directive (EWTD) on physicians.

Design: A systematic literature search was conducted in PubMed and Embase. Study selection, quality appraisal and data extraction were carried out by independent pairs of researchers using pre-established criteria.

Setting: Physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.

Participants: The total number of participants was 14,338.

Primary and secondary outcome measures: Health effects classified under the International Classification of Diseases (ICD-10).

Results: Over 3,000 citations and 110 full articles were reviewed. From these, eleven studies of high or intermediate quality carried out in North America, Europe and Japan met the inclusion criteria. Six studies included medical residents, junior doctors or house officers and the five others included medical specialists or consultants, medical, dental and general practitioners and hospital physicians. Evidence of an association was found between percutaneous injuries and road traffic accidents with extended hours/day or very long working weeks (LWH). The evidence was insufficient for mood disorders and general health. No studies on other health outcomes were identified.

Conclusions: LWH could increase the risk of percutaneous injuries and road traffic accidents, and possibly other incidents at work through the same pathway. While associations are clear, the existing evidence does not allow for an established causal or "dose-response" relationship between LWH and incidents at work, or for a threshold number of extended hours above which there is a significantly higher risk and the hours physicians could work and remain safe and healthy. Policy makers should consider safety issues when working on relaxing EWTD for doctors.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- The first systematic review, to our knowledge, on the effects of Long Working Hours (LWH) on physicians' health and safety.
- A systematic literature search conducted in PubMed and Embase with over 3,000 citations and 110 full articles reviewed.
- Eleven studies of high or intermediate quality carried out in North America, Europe and Japan, involving physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.
- The findings of this review confirmed that long working hours are associated with both an increased risk of needlestick injuries and road traffic accidents. Evidence was assessed as low or insufficient for mood disorders and general health.
- This review uncovered the lack of literature on the effect of long working hours on the general health of doctors, a topic that has important health ramifications not only for physicians but for patients too.

KEY WORDS: Working time, "Physicians"[MeSH], "Morbidity"[MeSH], "Occupational Injuries"[MeSH], Systematic review.

INTRODUCTION

The European Working Time Directive (EWTD),[1] applicable to all occupations across the EU, requires a maximum working week of 48 hours and establishes rest periods.[2, 3] Since it came into force in healthcare in 2009, it has been associated with concerns about the provision of health services including continuity of care, lower staffing levels, introduction of shift working, a reduction in training time and the adequate supervision of junior doctors.[4, 5] The large inter-country variations in physicians working hours as well as the mandatory controls over work hours across occupations raises further questions as to the basis of selecting and setting these limits and restrictions.[6] There is evidence of variable compliance to the EWTD within health care across Europe.[4, 5] The medical profession is also increasingly feminised with potential implications for women of childbearing age, which may coincide with the time in the profession where long working hours (LWH) are more likely to happen (i.e. during residency). Constraint on public-sector finances and increasing health care demands, have stimulated the debate on physicians' working hours,[7] with some evidence showing that the quality of health care is positively correlated with the number of physicians available to deliver that care.[8]

The Union of European Medical Specialists (UEMS) was created in 1958 to represent medical specialists in the EU, promote a high standard of training and practice at European level and facilitate the free movement of physicians across European countries.[9] The Occupational Medicine section was created in 1997 and is involved in issues related to specialist training, professional practice and research;[10] elaborating and publishing reference documents,[11] conducting and contributing to surveys,[12-14] and working in partnership with other key European groups.

The trigger for this project was a request by the Council of the UEMS to its Occupational Medicine Section for a position statement on the EWTD and possible health consequences to physicians of a potential removal of this professional group from the current EWTD 48 hours per week limit.[15-17] An evidence based approach was taken by the Section and a systematic review of the literature was undertaken.[15]

LWH have been associated with several adverse effects on workers' health, such as hypertension, cardiovascular disease, stress, depression, musculoskeletal disorders, chronic infections, diabetes, general health complaints and all-cause mortality in a variety of occupational groups.[18-26] In physicians, there is some evidence suggesting

that LWH can result in impacts such as stress, depression, burnout, injuries, fatigue and sleep deprivation,[27-29] and overwork has been considered one of the most stressful features of physicians' work.[27] Some evidence exists of the beneficial effect of reducing the number of working hours on the quality of life of medical residents [30] and on burnout.[27]

However, the relationship between work and health is complex. There is a substantial body of evidence showing that worklessness is associated with poorer health, whereas work is generally good for health and well-being,[31, 32] provided that the work environment is reasonably acceptable and supportive.[33]

Occupational and non-occupational exposures may play an important role such as the pattern and distribution of working hours, breaks, and recovery periods. Also, psychosocial and organisational factors, such as workload, job control, managers and peer support, training opportunities, individual characteristics and attitudes may be important.[6, 34]

Finally, physicians are a highly qualified, devoted and motivated professional group. Despite other existing reviews of a more general occupational approach, a review focused on physicians is warranted.

The aim of this project was to systematically review whether long-working hours (LWH) -i.e. defined as more than the 48 hour per week limit imposed by the EWTD- are associated with health effects –i.e. classified under the International Classification of Diseases (ICD-10) [35]- in physicians, and to examine what these associations are. This review does not include the potential impact of LWH on patient care or physician training.

METHODS

The systematic review sought to respond to the following research question: are LWH associated with health effects in physicians? For the purpose of this review, exposure to LWH was defined as working for more than 48 hours per week.[1] Health outcomes included any disease as defined in the ICD-10 [35] and work related injuries. If self-assessed tools on health status were used, only those validated were considered. Burnout, stress and outcomes such as individual symptoms, signs or biological markers (blood pressure, ECG, etc.) were not considered. Physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.

Study identification

Electronic searches were carried out using PubMed and EMBASE as search engines (December 2011). Our search strategy was similar in both databases and consisted of a combination of the following keywords and Mesh terms: night shift, morning shift, evening shift, afternoon shift, shift work, rotating shift, shift combination, shift duration or length, shift system, clockwise rotation, shift roster, extended shifts, night work, evening work, work schedule, work hours, starting time, early start, irregular working hours, direction of rotation, overwork, extended hours, shift rota, workload, work schedule tolerance, sleep deprivation, sleep disorders, chronobiology disorders, circadian rhythm, psychomotor performance, circadian disruption, vigilance, alertness, wakefulness, drowsiness, fatigue, insomnia, hypersomnolence, dyssomnia, eveningness, morningness, neurocognitive performance, concentration difficulties, arousal, health, morbidity, mortality, disease, illness, stress, strain, distress, accident, injur*, death, suicid*, education, medical, physician, medical staff, hospital, doctor, surgeon, house officers, medical school, surgery, surgical.

In addition, the references lists of articles selected for inclusion were carefully reviewed to identify additional studies.

Study selection

Studies were included if they measured the association between the exposure to LWH and health effects in physicians, used an observational epidemiological (i.e. cohort, case-control, cross-sectional) or experimental design. Articles relating to on-call, night or shift work, but with unknown exposure or exposed to less than 48 working hours per week were excluded. They were also excluded if the working groups were other than physicians or addressed only other exposures or outcomes (i.e. patient safety, fatigue, sleep deprivation, social or family disruption). No limitation was set for languages.

A total of 2,036 citations were obtained from PubMed and 980 non-duplicated citations using Embase, yielding a total of 3,016 citations that were all screened by independent pairs of reviewers. All titles and, where necessary, abstracts were reviewed. Ninety two potentially suitable publications were identified from the electronic searches, and 18 further studies were identified from the references lists, yielding a total of 110 studies published in English, French, German, Italian, Slovenian and Spanish. For all of them the full text was obtained and reviewed by independent pairs of reviewers. Disagreements within pairs were resolved by discussion and, where necessary, by a third reviewer who made the final decision. All authors participated as reviewers for screening the citations and full papers using well defined and pre-established criteria. Ninety nine articles were excluded at this stage. One further study was identified that was published after the search period. Finally, 12 papers were considered for quality assessment (Figure 1).

The methodological quality of the 12 studies was assessed by independent pairs of reviewers. A standardized 16 point scale based on CONSORT and STROBE statements and adapted from a previous systematic review [36] was used. It includes 16 items grouped into 6 areas: 1) objectives, 2) study design, 3) target population and sample, 4) variables, 5) data sources, collection and measurement, and 6) statistical methods. Each item was rated as 1 (the requirement was met), 0.5 (the requirement was partially met) or 0 (the requirement was not or unclearly met). Disagreements within pairs of reviewers were resolved by consensus or, where necessary, by a third reviewer. For each of the 12 studies, a final score based on the sum of all items was assigned and the percentage was calculated based on a maximum score of 16. Study quality was rated as low, moderate or high if it scored less than 60%, between 60% and 79.9%, and 80% or more of the maximum score, respectively.

One article of low quality was excluded [37] and a total of 11 studies of intermediate or high methodological quality were included for the purpose of this review (figure 1).

Data extraction

Selected information was extracted from each paper, including publication year, country of origin, study design, setting, study population, sample size, response rate, measure of working hours, definition used for LWH, health outcomes and their measurement, and main results on point risk estimates or frequencies of health outcomes, and their corresponding 95% confidence interval (95%CI), and whether the analyses were adjusted for potential confounders.

Evidence synthesis

To summarise the results on the relationship between LWH and health effects, levels of evidence synthesis was performed. This was based on the methodological quality, study design and the consistency of the study outcomes. The following criteria were based on two previous relevant systematic reviews.[38, 39]

Strong evidence: consistent results in more than two studies of high quality.

Moderate evidence: consistent results in one high quality study and one intermediate, or between some studies of intermediate quality.

Insufficient evidence: identification of only one study or inconsistent results across studies;

Evidence of no association: consistent results of a non-association in two or more studies.

A measure of the possible magnitude of the association was attempted using the following criteria based on the association point estimate (RR = relative risk, rate ratio or odds ratio) when available or otherwise the comparison between two frequencies was used:

- high: RR > 3·00
- intermediate: RR = 1·50-3·00
- low: RR = 1·01-1·49
- no association: RR=1·00

- unclear

RESULTS

We identified 11 studies related to LWH and health effects in physicians that were eligible for inclusion. The outcome of the methodological quality assessment is given in table 1.

Methodological quality was appraised as high for 3 of the 11 studies, with scores ranging from 84 to 94%. The other 8 studies were considered as intermediate quality as their scores ranged from 63 to 78%. The majority of all included studies received positive scores on items describing the study objectives, design and the study population (items from areas 1, 2 and 3), as well as the description of exposure and outcome variables (item 4). However, although data sources and collection were well described, reliability and validity of exposure to LWH and health outcome measures had low scores (items 5b and 5c). The statistical methods were in general appropriate, however the confounding variables were not measured in several studies and, as can be seen, the studies did not describe how they addressed missing data (items 6a and 6b).

Study characteristics:

The characteristics of the included studies are described in table 2.

Table 1. Methodological quality appraisal of the included studies

Study ID			Ayas 2006 [40]	Barger 2005 [41]	Firth-Cozens 1987 [42]	Fisman 2007 [43]	Hayasaka 2007 [44]	Kirkcaldy 1997 [45]	Rosta 2011 [46]	Stamp 2005 [47]	Sundquist 2000 [48]	Varma 2012 [49]	Zahrai 2011 [50]
1	Objectives	Are the objectives or hypotheses of the research described in the paper?	1	1	1	1	1	1	1	1	1	1	1
2	Study design	Is the study design presented?	1	1	1	1	1	1	1	1	1	1	1
3a	Target population	Do the authors describe the target population they wanted to research?	1	1	1	1	1	1	1	1	1	1	1
3b	Sample	Was a random sample of the target population taken? AND was the response rate 60% or more?	0.5	0.5	1	0.5	0.5	0.5	0.5	1	1	1	0.5
3c	Sample	Is participant selection described?	1	1	1	1	0.5	1	0	1	1	1	1
3d	Sample	Is participant recruitment described, or referred to?	1	1	1	1	1	1	1	1	1	1	0
3e	Sample	Are the inclusion and/or exclusion criteria stated?	1	1	1	1	0.5	0.5	1	1	1	1	0.5
3f	Sample	Is the study sample described? (minimum description: sample size, gender, age and occupation)	1	1	0	1	1	1	0	0	0	1	1
3g	Sample	Is the number of participants at each stage of the study reported?	1	1	1	0.5	1	0.5	1	1	1	1	1
4	Variables	Are the measures of long working hours and the health outcome described?	1	1	0.5	1	1	1	1	1	1	1	1
5a	Data sources, collection	Do authors describe the source of their data (e.g., official registry, health survey) AND how the data were collected?	1	1	1	1	1	1	1	1	1	1	1
5b	Measurement	Was reliability of the measure(s) of long working hours mentioned or referred to?	0	0	0	1	0	0	0	1	0	1	0
5c	Measurement	Was the validity of the measure(s) of long working hours mentioned or referred to?	1	1	0	0	0	0	0	0	0	0	0
5d	Measurement	Were health outcomes assessed by objective measures or validated self-reporting instruments?	0	1	1	1	1	0	1	1	1	1	1

6a	Statistical methods	Were appropriate statistical methods used and described, including those for addressing confounders?	0.5	1	0.5	1	0.5	0.5	1	0.5	1	1	0.5
6b	Statistical methods	Were the numbers/% of participants with missing data for long working hours and the health outcome indicated AND If more than 20% of data in the primary analyses were missing, were methods used to address missing data?	0	0	0	0.5	0	0	0	0	0	1	0.5
Total score** (%)			12.0 (75)	13.5 (84)	11.0 (69)	13.5 (84)	11.0 (69)	10.0 (63)	10.5 (66)	12.5 (78)	12.0 (75)	15.0 (94)	11.0 (69)
Quality rate***			interm	high	interm	high	interm	interm	interm	interm	interm	high	interm

** Maximum score = 16; *** Quality rate (%): low=<60; intermediate (interm)=60-79; high=80-100.

Table 2. Characteristics of the included studies.

Study ID	Country Of Study Population	Design*	Setting	Participants and sample size (response rate)	Working hours	Health outcomes	Quality score (%)
Ayas 2006 [40]	United States	Cohort, prospective (1 year) / Case-crossover	Hospital	Interns in postgraduate residency programs Sample size: 2,737 (80%)	Mean (SD) hrs worked/month: 249.8 (75.3). Self-reported, monthly survey. Strong correlation (Pearson *r*=0.76; *p* 0.001) with hours worked 244 (69.3) from work diaries of randomly selected subset of 192 interns. Extended periods (20 or more consecutive hours) vs. non extended periods (12 hrs or less consecutive hours).	Self-reported percutaneous injuries	12.0 (75)
Barger 2005 [41]	United States	Cohort, prospective (1 year)	Hospital	Residents first postgraduate year (interns). Sample size: 2,737 (80%)	Mean (SD) hrs worked/week 70.7 (26.0). Extended shift (≥ 24hrs) vs. non extended shift (<24 hours). Self-reported, validated.	Self-reported and documented motor vehicle crashes.	13,5 (84)
Firth Cozens 1987 [42]	United Kingdom	Cohort, prospective (1 year)	Hospital	Junior house officers Sample size: 170 (72%)	Mean number of hours per week = 90.6 hours (include on call). Self-reported.	GHQ-12 (case: score ≥2) and SCLDS, self-reported	11.0 (69)
Fisman 2007 [43]	United States and Canada	Case-Crossover	Hospital	Medical trainees vs. other HCW. Sample size: 109 vs. 241 (46%)	Median number of hours per week: medical trainees = 70; other HCW = 40 (p<0.001). Self-reported, high reliability.	Reported percutaneous injuries to employee health care service.	13.5 (84)
Hayasaka 2007 [44]	Japan	Cross-sectional	Hospital, clinics, other	Female physicians Sample size: 367 (63%)	Comparison of increasing number of hrs/week from ≤ 30 to >50, self-reported.	GHQ-30 (case: score ≥8), self-reported	11.0 (69)
Kirkcaldy 1997[45]	Germany	Cross-sectional	Not specified	Medical and dental practitioners. Sample size: 2,500 (not specified)	Long hours: mean 58.36 hrs/week SD 9.16 Short hours: mean 38.17 hrs/week SD 7.72 Self-reported.	Self-reported traffic accidents	10.0 (63)
Rosta 2011[46]	Germany and Norway	Cross-sectional	Hospital	Hospital physicians Participants: 1,917 (58%) in Germany and 1,072 (65%) in Norway answered the questionnaire. From survey responders, 1,260 (65.7%) Germans and 562 (52.4) Norwegians were included in the analysis after applying further inclusion criteria.	German vs. Norwegian physicians (%): Hours per day >9 = 58.8 vs. 26.7 60 hours on-call per month = 63.4 vs. 18.3 Self-reported.	Validated questionnaire on self-rated health	10.5 (66)
Stamp 2005 [47]	United States	Before-and-after	Hospital	Residents of general surgery Sample size: 28 (97%)	Changes of work patterns, after reduction from 90-110 to 78-80 work hrs/week.	SF-36, mental health; BDI, depression; self-reported	12.5 (78)

Sundquist 2000 [48]	Sweden	Cross-sectional	Primary care	General practitioners Sample size: 1,004 (72%)	Overtime defined as working at least 47 hrs/week. Self-reported.	Swedish SF-36, impaired mental health, self-reported	12.0 (75)
Varma 2012 [49]	Denmark	Cohort, prospective (20 months)	Hospital	Senior medical consultants Sample size: 2,790 (62%)	Long work hours (>40hrs/week). Self-reported.	Depression: redemption of anti-depressive drug prescriptions	15.0 (94)
Zahrai 2011 [50]	Canada	Cohort, prospective (6 months)	Hospital	Orthopaedic surgery residents Sample size: 16 (not specified)	Night float (n=9): 77.8% did >80h/week at baseline; 71.4% at follow-up). Standard call (n=7): 57.1% did >80h week at baseline; 80% at follow-up.	SF-36, mental health score, self-reported	11.0 (69)

* Follow-up period in brackets for prospective cohort studies; GHQ-12= General Health Questionnaire-12 items; SCLDS= Symptom Checklist Depression Scale; BDI = Beck Depression Inventory II.

Five had been carried out in North America (4 in the United States [40, 41, 43, 47] and 1 in Canada [50]), five in European countries (Denmark,[49] Germany,[45] Norway and Germany [43]; Sweden,[48] and the United Kingdom [42]) and one study in Japan.[44] Four used a prospective cohort design (follow-up ranging from 6 to 20 months), 2 were case-crossover, 4 cross-sectional, and 1 was a before and after study. Overall, the total number of participants was 14,338 and included medical residents, junior doctors or house officers only working in hospitals, medical specialists or consultants, medical and dental practitioners, general practitioners, hospital physicians and one study included female physicians from a community service. Most studies (n=9) compared different working hour patterns within the same group of physicians, one compared physicians from two different countries,[46] and one included other health care workers as the comparison group.[43]

LWH was defined as more than 48hours/week in most studies, and some did not provide a definition. Two studies considered LWH below 48/week.[46, 48] All the included studies investigated health effects associated with working more than 48h/week, with number of hours ranging between <30 to 110hours/week. One paper studied the effects of an increasing number of hours from 30 to more than 50h/week.[44] Others studied more than 40h/week,[49] 47h/week,[48] above 58h/week [46](two studies) and 6 studies referred to more than 70h/week,[40-43, 47, 50] one of which compared 78-80h/week to 90-110 weekly hours.[47] Information on working hours was self-reported in nine studies, with reasonably good validity only in two studies [40, 41] and reliability in one.[43] In the other two studies it was obtained from established work schedules.[47, 50]

Identified health outcomes that fulfilled the inclusion criteria were percutaneous injuries, motor vehicle accidents, mood disorders and general health. Information was self-reported in all studies using well known validated questionnaires, except three studies that used documented information on motor vehicle crashes,[40] reported percutaneous injuries [42] and antidepressant prescription data as a surrogate of depression.[49]

Summary of findings:

Table 3 shows the findings from each of the 11 studies.

Table 3. Results on the association between long working hours and identified health effects in physicians.

Health outcome	Main results	Adjustment by confounders	Study ID
Percutaneous injuries	Odds ratio (95% CI) for injuries during extended (>20 hours/day) vs. non extended periods (first 12 hours at work): All percutaneous injuries: 1.61 (1.46-1.78); Injuries reported to OH: 1.83 (1.48-2.28); Injuries in the ICU: 1.87 (0.69-5.04); Injuries in the operating room or labour and delivery: 1.77 (1.49-2.09); Injuries in the ICU, non-ICU, or ED: 2.17 (1.56-3.00)	Time of day and circadian influences	**Ayas 2006**[40]
Percutaneous injuries	Medical trainees vs. other health care workers: total median working hours: 70 vs. 40 hours (p<.001); median previous working hours at the time of injury: 6.5 vs. 5 hours (p< .001). Odds ratio (95% CI) for self-reported fatigue: work > 40h in the past week: 3.50 (2.06-5.92); work during more than 5d in past week: 4.20 (4.46-7.15); >12h at work before injury: 8.58 (3.7-19.86). Incidence rate ratio (95%CI) for association between fatigue and injuries: all workers: 1.40 (1.03-1.90); medical trainees: 2.94 (1.71-5.07); other HCWs: 0.97 (0.66-1.42) (p=0.001 for heterogeneity).	Age and sex	**Fisman 2007**[43]
Motor vehicle accidents	Odds ratios (95% CI), after extended shift (≥ 24hrs) vs. non extended shift (<24 hours): Crashes = 2.3 (1.6-3.3); Near miss accidents = 5.9 (5.4–6.3)	Age and sex	**Barger 2005**[41]
Motor vehicle accidents	Incidence rates of driving accidents on house visits: Long hours (mean 58.36 hrs/week) = 0.10 (0.31) vs. short hours (mean 38.17 hrs/week) = 0.02 (0.18). p<0.05	Not adjusted	**Kirkcaldy 1997**[45]
Mood disorders, depression	Hazard ratio (95% CI) hours/week intervals and redemption of anti-depressive drug prescription(reference group: 37-40 h/week): 25–36 h: 0.83 (0.24–2.82); 41–44 h: 0.95 (0.50–1.77); 45–49 h: 0.88 (0.43–1.78); 50–54 h: 0.83 (0.32–2.14); 55–59 h: 0.67 (0.15–2.94); >60 h: 0.48 (0.06–3.68). Cox regression analysis of work hours as a continuous variable: HR=0.93 (0.76–1.13)	Age, sex, marital status, medical specialty, decision authority at work, social support at work, quantitative work demands and previous redemption of AD drug prescription	**Varma 2012**[49]
Mood disorders, GHQ-30	Prevalence of cases by working time (hours/week) ≤ 30 hours: 35.7% (41/115); >30 to 40 hours: 39.0% (39/100); >40 to 50 hours: 37.7% (26/69); >50 hours: 56.8% (46/81); *p*=0.0179 Stepwise multiple logistic regression, >50 h/wk. vs. ≤50 h/wk.: parameter 0.635 (p=0.0293)	Marital status, medical facility, position, and night duty.	**Hayasaka 2007**[44]
Mood disorders, GHQ-12 and SCLDS	No association was found between number of hours worked in a week and depression.	Not adjusted	**Firth Cozens 1987**[42]
Mood disorders, SF-36	Scores when night float vs. standard call at baseline and follow-up: Mental health mean score (SD): baseline = 57.33 (22.63) vs. 65.71 (7.61); follow up = 52.00 (15.49) vs. 60.80 (11.45); p=0.72 Mental health component summary (SD), baseline = 34.84 (14.06) vs. 40.21 (7.61) ; follow up = 30.15 (10.71) vs. 42.40 (6.23); p=0.39 Regression analysis: increased number of hours in hospital correlated with significantly lower SF-36 scores in almost all domains.	Not adjusted	**Zahrai 2011**[50]
Mood disorders, SF-36 and BDI	Improvement of levels before and after duty hours limitations from 90-110 to 78-80 hors/week): Mental health SF-36): no statistically significant improvement Individual questions BDI before and after duty hours limitations, only energy level statistically significant.	Not adjusted	**Stamp 2005**[47]

Health outcome	Main results	Adjustment by confounders	Study ID
Mood disorders, SF-36	Odds ratio (95% CI)for working 47 hours per week or more: Impaired mental health, men: 1.59 (0.95-2.66); women: 1.86 (1.03-3.37)	Age and amount of time in practice	**Sundquist 2000**[48]
General health, SF-36	SF-36: Scores when night float vs. standard call at baseline and follow-up: General health mean score (SD): baseline = 62.11 (17.47) vs. 77.57 (24.25); follow up = 56.43 (24.89) vs. 84.20 (16.50); p=0.41 Physical health component summary (SD): baseline = 46.16 (13.15) vs. 52.01 (13.33); follow up = 39.32 (9.80) vs. 56.15 (2.18); p=0.015 Vitality mean score (SD): baseline = 51.67 (14.58) vs. 51.43 (15.74); follow up = 48.57 (14.92) vs. 51.00 (10.84); p=0.20 Regression analysis demonstrated that the increased number of hours spent in hospital correlated with significantly lower general health, physical function, mental health, role emotional, social function and mental component summary scale scores (all p < 0.05)	Not adjusted	**Zahrai 2011**[50]
General health, SF-36	SF-36, BDI: Improvement of levels before and after duty hours limitations from 90-110 to 78-80 hours/week): Physical health: no statistically significant differences.	Not adjusted	**Stamp 2005**[47]
General health, SF-36	SF-36: Odds ratio (95% CI) adjusted for age and time in practice, for working 47 hours per week or more: Impaired general health: men = 1.66 (1.00-2.77); women = 1.59 (1.00-3.17)	Age and amount of time in practice	**Sundquist 2000**[48]
General health	Self-rated health: Odds ratio (95% CI) of good self-rated health (logistic regression): Norwegian work time pattern**= 1.35 (1.03-1.77); working in Norway = 4.17 (3.02 - 5.73).	Age, sex and country of work	**Rosta 2011**[46]

**Not working more than 9 hours a day and having more than 60 hours a month on-call); GHQ-12= General Health Questionnaire-12 items; SCLDS= Symptom Checklist Depression Scale; BDI = Beck Depression Inventory II.

Percutaneous injuries: One study of high quality and one of intermediate quality, with non-adjusted analyses, showed consistent results in medical trainees. Ayas et al. found an increased risk of percutaneous injuries associated with working more than 20 hours/day compared to non-extended hour periods, except in intensive care units;[40] Fisman et al. found an association between self-reported fatigue and the number of working hours at the time of injury and a 3-fold increased risk of sharp injuries in medical trainees compared to other health care workers (median of working hours/week before the injury: 70 vs. 40;p<.001).[43]

Motor vehicle accidents: Two studies of high and intermediate quality showed that long working weeks were associated with an at least two-fold increased risk. Barger et al. found adjusted increased risks for car crashes (OR 2·3; 95%CI 1·6-3·3) and near miss accidents (OR 5·9; 95%CI 5·4-6·3) associated with working extended shifts.[41] Kirkcaldy et al. showed that non-adjusted incidence rates of traffic accidents on house visits was five times as much when physicians worked for a mean of 58 hours/week compared to 38 hours/week (p<0,05).[45]

Mood disorders: Six studies of intermediate and high quality found contradictory results for mood disorders. Three of them, all of intermediate quality, provided evidence of an association between LWH and mental health problems. The prevalence of GHQ-30 cases was higher for female physicians working above 50 hours/week than for those working 30 or less hours/week (p<0.05),[44] and an increased adjusted risk of impaired mental health was found for men and women general practitioners who worked for 47 hours per week or more.[48] Zahrai et al. found that increased number of hours spent by residents in hospital correlated significantly with lower mental health and mental component summary scale scores (SF-36).[50] Three other studies, one of which of high quality, did not find an association between depression or mood disorders and LWH. Varma et al. did not find higher adjusted risks of depression associated with increased number of hours in senior medical consultants, even when they worked for more than 60 hours per week.[49] Depression or mood disorders measured with GHQ-12, SCLDS or BDI, all validated tools for mental health, were not associated with increased number of working hours [42] nor improved after decreasing the number of hours from 90-110 to 78-80 hours per week in residents.[47]

General health: Four studies of intermediate quality analysed the association between LWH and general health. Zahrai et al. found that increased number of hours spent by

residents in hospital correlated significantly with poorer general health, physical function and vitality using SF-36.[50] A comparative study of physicians in two different countries found that Norwegians showed higher non-adjusted prevalence of self-reported good health compared to physicians in Germany who worked longer hours.[46] An increased adjusted risk of impaired general health was found for men and women general practitioners who worked for more than 47 hours per week.[48] However, following the implementation of reducing the number of hours from an average of 90-110hours/week to 78-80hours/week did not lead to an overall improvement of residents self-reported physical health.[47]

Synthesis of the evidence:

The levels of evidence synthesis obtained from the analysis of the 11 studies included in this review are shown in table 4.

Table 4. Available scientific evidence on the health problems associated with long working hours in physicians: levels of evidence synthesis.

Health outcome	Degree of evidence*	Magnitude of the association**	Studies ID
Percutaneous injuries	++	++	Ayas 2006[40], Fisman 2007[43]
Motor vehicle accidents	++	++	Barger 2005[41], Kirkcaldy 1997[45]
Mood disorders	+	+/-	Varma 2012[49], Hayasaka 2007[44], Firth Cozens 2001[42], Zahrai 2011[50], Stamp 2005[47], Sundquist 2000[48]
General health	+	+/-	Zahrai 2011[50], Stamp 2005[47], Sundquist 2000[48], Rosta 2011[46]

* Strong evidence (+++): consistent results in more than 2 studies of high quality; Moderate evidence (++): consistent results in two studies of high quality, or one high quality study and one intermediate, or between some studies of intermediate quality; Insufficient evidence (+): identification of only one study or inconsistent results across studies; Evidence of no association (-): consistent results of a non-association in two or more studies.

** Magnitude of the association: it refers to the magnitude of the association point estimate (RR = relative risk, rate ratio or odds ratio): high (+++) if RR > 3; intermediate (++) if RR = 1·5-3; low (+) if RR < 1·5; no association (-); unclear (+/-).

There is moderate evidence of an association between LWH and percutaneous and motor vehicle accidents in physicians. This evidence comes from one study of high quality and another of intermediate quality both for percutaneous injuries, and for motor vehicle accidents and the magnitude of this association could be intermediate.

For mood disorders and general health there is a low or insufficient level of evidence of an association with LWH in physicians. This comes from 6 studies on mood disorders and 4 on general health of high and intermediate methodological quality and inconsistent results among them. No conclusion can be drawn about the magnitude of such associations if existed.

DISCUSSION

This review found moderate scientific evidence for a positive association of intermediate magnitude between exposure to LWH and percutaneous injuries and motor vehicle accidents in physicians. Evidence was assessed as low or insufficient for mood disorders and general health.

To our knowledge, this is the first systematic review on the effects of LWH on physicians' health and safety. The extensive searches were restricted to indexed journals. We systematically identified, selected and assessed the methodological quality of studies by means of independent pairs of reviewers. The quality assessment form and the levels of evidence used in this review were based on CONSORT and STROBE statements and on previous systematic reviews.[36, 38, 39] To give some estimate of the potential magnitude of the effect, we added levels based on the point estimates of measures of association provided by the studies. However, such levels had not been previously established.

Strengths and limitations of the studies

Four studies had a longitudinal prospective design, however only two of them used a robust methodology. Two other studies used a crossover design, which reduces the likelihood of individual variability and confounding, and were considered well designed. Four studies, all rated as intermediate methodological quality, were cross sectional in design and therefore it is difficult to draw causal relationships from them.

Although reverse causality cannot be ruled out from cross-sectional designs, it is unlikely that poorer health determines longer working hours than their healthier peers. One study used a quasi-experimental design but lacked a control group. Also, because of different designs used across the included studies, different estimates of the effects were used that may not be directly comparable.

Other limitations include the variability, validity and reliability of working hours, which in most studies was self-reported. Although only those papers specifying the number of working hours were accepted, shift and night work might have worked as confounders too, as it is difficult to acknowledge which proportion of the health effects observed was due exclusively to the LWH component and not due, for instance, to stress, fatigue or sleep deprivation. Moreover, the studies did not discriminate between time spent actively working or asleep while on call, though the EWTD considers all hours on call as working time. Likewise, it is possible that participants were aware that the study was conducted to assess associations between the number of hours worked and health outcomes. Therefore, there might be a volunteer bias in those doctors who work extended hours. To minimise this bias, a case-cross over analysis has been performed in three of the included study [40, 41, 43].

Health outcomes were measured mainly based upon self-report, with the exception of documented motor vehicle crashes,[40] reported percutaneous injuries [42] and register-notified anti-depressive drug prescriptions.[48] No studies with other objective measures of health (e.g., mortality) or mental disease (e.g., hospital data) were identified. However, self-reported or perceived health was assessed by validated and widely used instruments, especially when health status (general, mental or physical health) or ill-health symptoms are evaluated. Nonetheless, this raises the possibility that the observed associations might reflect differences in propensity to report health problems when they occur rather than true differences in the risk of worse health status. We cannot rule this out, but the higher propensity for reporting among those who work for longer hours and poorer work schedule patterns would have to relate specifically to the reporting of ill-health rather than non-ill health. It seems unlikely that major differences in propensity to report would extend to a more concrete outcome.

The majority of the studies of this review took into account confounding variables in their analyses, such as sex and age. However, none of the studies analysed the potential effect of modifying factors, such as psychosocial aspects at work, including attitudes,

motivation, job requirements, demands and content, organisational climate, social relationships at work, work satisfaction, supportive organisation, or the relative number of physicians and other staff available for patient care. Neither were important aspects outside of work, such as life events or lifestyles, taken into account. The study of Japanese female physicians found significant differences by marital status with higher GHQ scores for those who were married.[44] There are significant differences in the number of physicians per capita across different countries and this together with the structure and organisation of health services are potential confounders too that were not addressed in the selected studies across countries.

A relevant outcome from this review would be to establish a threshold number of extended hours above which there is a significantly higher risk. There is a lack of evidence for a dose-response relationship and nor does the evidence give any indication for a threshold number of hours that physicians could work and remain safe and healthy.

Finally, the heterogeneity of the included studies did not allow a more quantitative synthesis, and a level of evidence approach was used instead.

Comparison with other studies

Our findings of an association between LWH and injuries and accidents in physicians are consistent with research in other occupational groups. Dembe et al. carried out a longitudinal survey in the US including more than 10,000 workers from a variety of occupations and settings.[19] A strong dose-response effect (adjusted for age, gender, occupation, industry and region) was found between LWH and injuries above 40h/week and 8h/day; jobs with overtime schedules were associated with a 61% higher injury hazard rate compared to jobs without overtime; working at least 12 hours per day was associated with a 37% increased hazard rate, and working at least 60 hours per week was associated with a 23% increased hazard rate. They concluded that LWH might indirectly precipitate workplace injuries through a causal process by inducing fatigue or stress in workers. In a separate study assessing injury risks to health care personnel, Dembe et al. demonstrated that the risk of injury when working overtime or at least 60 h per week among physicians and nurses was statistically significant.[51] In our review we found that information about the length of shift work varied across studies, some

analysing shifts of up to twenty four hours. Work that included such shifts was more strongly associated with accidents and injuries.[40, 41]

We found insufficient evidence of an association between LWH and mood disorders, other diseases and general health in physicians, despite some scientific evidence suggesting that LWH increase morbidity and mortality in other occupations.[19-26]

Possible explanations for these inconsistencies might be the scarcity of high quality research, and the fact that some of the included papers did not fully address the effect of confounding factors, and the possibility that working as a physician might have a protective effect by itself, for instance through higher motivation and social recognition relative to other occupations.

Stamp et al., in the only before and after study, did not find significant differences for mood disorder or general health before and after implementing requirements to reduce working schedules to 80h/week in residents.[47] It is debatable, however, whether any conclusions can be drawn from the impact of reducing extremely high working hours (90-110hours/week) to very high working hours (78-80hours/week).

Although we searched for and identified literature describing associations between LWH and increased mortality in doctors,[52, 53] none of those studies specified the number of hours worked, and therefore were not included.

Recommendations for future research

Further research, including well designed prospective and retrospective longitudinal studies, is needed in this field to establish on scientific grounds what number of hours are safe, both to patients and physicians. Also studies on long term health effects should be considered.

This review was aimed to study exclusively possible detrimental health effects of LWH (more than 48 hours per week) to physicians. Other systematic reviews should be conducted addressing issues, such as the impact of working hours on doctor's well-being and quality of life, the quality of junior doctors training, the quality and continuity of care and impact on patient safety. Also, from a staffing perspective, the economic and social burden of sick leave and physicians leaving the profession are important issues that were beyond the scope of this review and that should be considered in future

research. Other established criteria [54] that working hours should be 'family friendly', promote gender equality, enhance productivity or facilitate worker choice and influence over working hours, were not addressed. Further research is needed in these other areas.

Conclusions

The findings of this review have confirmed that long working hours are associated with both an increased risk of needlestick injuries and road traffic accidents. The fact that these associations are labelled moderate in this review, should not preclude safety procedures and measures being assessed and reassessed to ensure maximum physician and patient safety. The reduction of needlestick injuries requires improved use of safe sharp devices, compliance with safe working procedures, training and regular audits of working practices, particularly for junior doctors who are at greatest risk of needlestick injuries, possibly due to inadequate training.[55, 56] The risk of road traffic accidents could be reduced by discouraging driving after very long shifts (more than 16 hours) and where there is sleep deprivation.[57] Transport should be provided by the employer in these situations. This is a pragmatic recommendation which recognises that long working hours have not been eliminated as a result of the EWTD.

This review uncovered the lack of literature on the effect of long working hours on the general health of doctors, a topic that has important health ramifications not only for physicians but for patients too. However, the paucity of available studies, which does not allow for a causal or dose-response relationship to be established, raises further questions about the evidence-base of the current 48 hour limit, especially as this is not enforced strictly and junior doctors, for example in the UK, may opt-out.[58] It may be relevant that the EWTD was not based entirely on science but also took into account, as does all EU legislation, the views of "social partners".[3]

It is likely that the mechanism for these increased health and safety risks is fatigue, and our finding of increased risk of accidents to staff, may also be associated with increased risk of clinical errors. Patient safety was not the purpose of this review but long working hours are well recognised to cause decrement in performance both in health care and other professions, where the performance of staff can be safety critical, such as airline pilots, and professional drivers.[59]

In addition the EWTD has been associated with real concerns about the adequacy of the training of doctors given their reduced exposure to patient care, and the lack of

experience of the patient journey, imposed by adherence to the EWTD.[4] This is potentially exacerbated at the same time by the constraints on the overall duration of training imposed by other EU directives.[60] Further reviews of the EWTD, with regard to this unique group of workers, need to take into account other social factors such as the impact on patient care. This is particularly the case when all health care systems are under strain because of burgeoning demands and limited resources.[4]

The findings of this systematic review lead to the suggestion of the following guidelines in the event of non-compliance with the EWTD: 1) long shifts should be avoided to protect both physicians' health and patient safety and no shift should be longer than 16 hours (and then exceptionally);[28] 2) Physicians should be discouraged from driving after long shifts to reduce the risk of motor accidents;[41, 45] 3) rigorous attention should be paid to reducing the risk of sharps injuries;[40, 43] 4) organisational aspects such as workload and job control, as well as the pattern and distribution of working hours, breaks and recovery periods should be carefully taken into account, in order to avoid fatigue and sleep deprivation that could lead to mistakes and accidents;[61, 62] 5) Physicians should work in supportive psychosocial environments, in teams, and with adequate training and supervision as their wellbeing is important also for the health of their patients.[61, 62]

Further research is required to establish any longer term effects such as on mortality and mental health, also the impact of the nature of the work organisation taking into account the psychosocial aspects of the physicians working and non-working lives, and to determine how many hours are safe for physicians to work.

ACKNOWLEDGMENTS

Leena Isotalo, Trial Search Coordinator at the Cochrane Occupational Safety and Health Review Group, designed and conducted the final search of the evidence using Medline and Embase databases.

FUNDING

The Section of Occupational Medicine of the UEMS provided limited support to the work of all authors (MCR, ED, SVP, KS, AS, PR, RH), with the exception of the senior authors (EBM, CS). The expenses of one meeting of the core group (MCR, SVP, ED, EBM, CS) were also funded by the Section. **Role of funding source:** The Section of Occupational Medicine of the UEMS acted in its role of promoting and contributing to the research in Occupational Medicine.

AUTHORS' CONTRIBUTIONS:

Mari Cruz Rodríguez-Jareño, Evangelia Demou, Sergio Vargas-Prada, Alenka Škerjanc, Kaveh A Sanati, Ewan B. Macdonald and Consol Serra conceived and submitted the design, carried out the acquisition of data, analysis and interpretation of data, drafted the article and revised it.

Pedro G. Reis and Ritva Helimäki-Aro contributed to acquisition of data, analysis and interpretation of data, drafted the article and revised it.

All authors gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

CONFLICT OF INTERESTS

None. No financial relationships with any organisations that might have an interest in the submitted work, and no other relationships or activities that could appear to have influenced the submitted work.

DATA SHARING STATEMENT

No additional data are available

ETHICS COMMITTEE APPROVAL was not sought as no patients were involved in this study.

REFERENCES

1 EUR-Lex. Directive 2003/88/EC of the European Parliament and of the Council of 4 November 2003 concerning certain aspects of the organisation of working time. In: Union OJotE, ed. Brussels: Official Journal of the European Union, 2003.

2 EPHA. 2012 EPHA Briefing on the European Working Time Directive (EWTD) http://www.epha.org/IMG/pdf/EPHA_Briefing_EWTD_Oct2012_FINAL.pdf; Access Year 2012.

3 European Commission. 2013 Working Conditions-Working Time Directive. http://ec.europa.eu/social/main.jsp?catId=706&langId=en&intPageId=205; Access Date December 2013.

4 House J. Calling time on doctors' working hours. Lancet 2009;**373**:2011-12.

5 The Lancet. Doctors' training and the European Working Time Directive. Lancet 2010;**375**:2121.

6 Leff D, Aziz O, Darzi A. Trucks, planes, and scalpels - Is there an evidence-based approach to surgeons' working hours? Arch Surg 2007;**142**:817-20.

7 Parthasarathy S. Sleep and the medical profession. Curr Opin Pulm Med 2005;**11**:507-12.

8 Dr. Foster Hospital Guide. 2011 www.drfosterhealth.co.uk; Access Date February 2011.

9 UEMS. 2012 The Union of European Medical Specialists. www.uems.net; Acess Year 2012.

10 Macdonald EB, Ritchie KA, Murray KJ, et al. Requirements for occupational medicine training in Europe: a Delphi study. Occup Environ Med 2000;**57**:98-105.

11 Macdonald E, Baransky B, Wilford J. Occupational Medicine in Europe: Scope and Cometencies. In: Health WECfEa, ed. Health, Environment and Safety in Enterprises Series n. 3. Bilthoven, The Netherlands, 2000.

12 Ballester M, Cordell N, Rodriguez Jareno MC, et al. A European survey of professional bodies representing occupational medicine specialists. Occup Med (Oxf) 2012;**62**:366-70.

13 Cashman C, Slovak A. The occupational medicine agenda: Routes and standards of specialization in occupational medicine in europe. Occup Med (Oxf) 2005;**55**:308-11.

14 De Schryver A, Claesen B, Meheus A, et al. European survey of hepatitis B vaccination policies for healthcare workers. Eur J Public Health 2011;**21**:338-43.

15 Rodriguez-Juareno M, Serra C, Demou E, et al. European Working Time Directive: systematic review for evidence based decision making. SUMMARY REPORT. In:

Medicine USoO, ed. http://www.uems-occupationalmedicine.org/sites/default/files/old_userfiles/File%203%20-%20EWTD%20UEMS%20REPORT%2005%2010%202011.pdf,: UEMS Section of Occupational Medicine, 2011.

16 UEMS. 2010 The Newsletter of European Medical Specialists. http://www.uems.eu/__data/assets/pdf_file/0014/1463/1466.pdf.

17 UEMS. Meetings of the Board & Council of UEMS- Report. In: European Union of Medical Specialists; http://www.uems.eu/__data/assets/pdf_file/0011/1235/UEMS_2011.37_-_report_-_UEMS_Council_6-8.10.2010_Napoli.pdf, ed. Napoli (Italy), 2011.

18 Bannai A, Tamakoshi A. The association between long working hours and health: A systematic review of epidemiological evidence. Scand J Work Environ Health 2013.

19 Dembe AE, Erickson JB, Delbos RG, et al. The impact of overtime and long work hours on occupational injuries and illnesses: new evidence from the United States. Occup Environ Med 2005;**62**:588-97.

20 Johnson JV, Lipscomb J. Long working hours, occupational health and the changing nature of work organization. Am J Ind Med 2006;**49**:921-9.

21 Kivimaeki M, Batty GD, Hamer M, et al. Using Additional Information on Working Hours to Predict Coronary Heart Disease. Ann Intern Med 2011;**154**.

22 Liu Y, Tanaka H, Fukuoka Heart Study G. Overtime work, insufficient sleep, and risk of non-fatal acute myocardial infarction in Japanese men. Occup Environ Med 2002;**59**:447-51.

23 Sokejima S, Kagamimori S. Working hours as a risk factor for acute myocardial infarction in Japan: case-control study. BMJ 1998;**317**:775-80.

24 Spurgeon A, Harrington JM, Cooper CL. Health and safety problems associated with long working hours: A review of the current position. Occup Environ Med 1997;**54**:367-75.

25 Virtanen M, Ferrie JE, Singh-Manoux A, et al. Long working hours and symptoms of anxiety and depression: a 5-year follow-up of the Whitehall II study. Psychol Med 2011;**41**:2485-94.

26 Virtanen M, Stansfeld SA, Fuhrer R, et al. Overtime work as a predictor of major depressive episode: a 5-year follow-up of the Whitehall II study. PloS One 2012;**7**:e30719.

27 Gopal R, Glasheen JJ, Miyoshi TJ, et al.. Burnout and internal medicine resident work-hour restrictions. Arch Intern Med 2005;**165**:2595-600.

28 Reed DA, Fletcher KE, Arora VM. Systematic Review: Association of Shift Length, Protected Sleep Time, and Night Float With Patient Care, Residents' Health, and Education. Ann Intern Med 2010;**153**:829-42.

29 Sargent MC, Sotile W, Sotile MO, et al.. Quality of Life During Orthopedic Training and Academic Practice Part 1: Orthopedic Surgery Residents and Faculty. J Bone Joint Surg Am 2009;**91A**:2395-405.

30 Fletcher KE, Underwood W, Davis SQ, et al. Effects of work hour reduction on residents' lives - A systematic review. JAMA 2005;**294**:1088-100.

31 Black C, Frost D. *Health at work - an independent review on sickness absence*. London: The Stationary Office, 2011.

32 Waddell G, Burton K. *Is work good for your health and wellbeing?* London: The Stationary Office, 2006.

33 Butterworth P, Leach LS, Strazdins L, et al. The psychosocial quality of work determines whether employment has benefits for mental health: results from a longitudinal national household panel survey. Occup Environ Med 2011;**68**:806-12.

34 Lockley SW, Landrigan CP, Barger LK, et al. When policy meets physiology - The challenge of reducing resident work hours. Clin Orthop 2006:116-27.

35 ICD-10. 2012 International Classification of Diseases. http://www.who.int/classifications/icd/en/; Access Year 2012.

36 van Uffelen JGZ, Wong J, Chau JY, et al. Occupational Sitting and Health Risks A Systematic Review. Am J Prev Med 2010;**39**:379-88.

37 Kirsling RA, Kochar MS, Chan CH. An evaluation of mood states among 1st-year residents. Psychol Rep 1989;**65**:355-66.

38 Bernard B. A critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity and low back. In: NIOSH, ed. Publication No. 9741. Cincinnati: National Institute for Occupational Safety and Health, 1997.

39 Steenstra IA, Verbeek JH, Heymans MW, et al. Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature. Occup Environ Med 2005;**62**:851-60.

40 Ayas NT, Barger LK, Cade BE, et al. Extended work duration and the risk of self-reported percutaneous injuries in interns. JAMA 2006;**296**:1055-62.

41 Barger LK, Cade BE, Ayas NT, et al. Extended work shifts and the risk of motor vehicle crashes among interns. N Engl J Med 2005;**352**:125-34.

42 Firth-Cozens J. Emotional Distress in Junior House Officers. BMJ 1987;**295**:533-36.

43 Fisman DN, Harris AD, Rubin M, et al. Fatigue increases the risk of injury from sharp devices in medical trainees: Results from a case-crossover study. Infect Control HospEpidemiology 2007;**28**:10-17.

44 Hayasaka Y, Nakamura K, Yamamoto M, et al. Work environment and mental health status assessed by the general health questionnaire in female Japanese doctors. Ind Health 2007;**45**:781-86.

45 Kirkcaldy BD, Trimpop R, Cooper CL. Working hours, job stress, work satisfaction, and accident rates among medical practitioners and allied personnel. Int J Stress Manag 1997;**4**:79-87.

46 Rosta J, Aasland OG. Work Hours and Self rated Health of Hospital Doctors in Norway and Germany. A comparative study on national samples. BMC Health Serv Res 2011;**11**.

47 Stamp T, Termuhlen P, Miller S, et al. Before and after resident work hour limitations: an objective assessment of the well-being of surgical residents. Curr Surg 2005;**62**:117-21.

48 Sundquist J, Johansson SE. High demand, low control, and impaired general health: working conditions in a sample of Swedish general practitioners. Scand J Pub Health 2000;**28**:123-31.

49 Varma A, Marott J, Stoltenberg C, et al. With long hours of work, might depression then lurk? A nationwide prospective follow-up study among Danish senior medical consultants. Scand J Work Environ Health 2012;**38**:418-26.

50 Zahrai A, Chahal J, Stojimirovic D, et al. Quality of life and educational benefit among orthopedic surgery residents: a prospective, multicentre comparison of the night float and the standard call systems. Can J Surg 2011;**54**:25-32.

51 Dembe AE, Delbos R, Erickson JB. Estimates of injury risks for healthcare personnel working night shifts and long hours. Quality & Safety in Health Care 2009;**18**:336-40.

52 Lindfors PM, Nurmi KE, Meretoja OA, et al. On-call stress among Finnish anaesthetists. Anaesthesia 2006;**61**:859-64.

53 Šelb J, Albreht T. Mortality rates of medical doctors in Slovenia in 1985 to 1999. Zdravniški vestnik 2000:I47-I48.

54 Lee S, McCann D, Messenger J. Working Time Around the World: Trends in working hours, laws and policies in a global comparative perspective. Geneva, Switzerland: International Labour Office, 2007.

55 Elder A, Paterson C. Sharps injuries in UK health care: a review of injury rates, viral transmission and potential efficacy of safety devices. Occup Med-Oxford 2006;**56**:566-74.

56 Naghavi SHR, Sanati KA. Needlestick injuries: Does left-handedness matter? Am J Infect Control 2009;**37**:341-41.

57 Connor J, Norton R, Ameratunga S, et al. Driver sleepiness and risk of serious injury to car occupants: population based case control study. BMJ 2002;**324**:1125-28A.

58 BMA. 2013 European Working Time Directive, http://bma.org.uk/practical-support-at-work/ewtd,. Access Date December 2013.

59 Wilson AM, Weston G. Application of Airline Pilots Hours to Junior Doctors. BMJ 1989;**299**:779-81.

60 The European Specialist Medical Qualifications Order 1995. 1995 Statutory Instruments, 1995 No. 3208, MEDICAL PROFESSION, The European Specialist Medical Qualifications Order 1995, http://legislation.data.gov.uk/uksi/1995/3208/made/data.htm?wrap=true,. Access Date December 2013.

61 Boorman S. NHS Health and Well-being. The Boorman Review. Leeds, 2009.

62 Williams S, Michie S, Pattani S. Improving the health of the NHS workforce. Report of the partnership on the health of the NHS Workforce. London: The Nuffield Trust, 1998.

Figure legend

Figure 1. Results of the search strategy, using search engines on PubMed and EMBASE (December 2011) and screening of references lists of identified full papers, study selection and quality assessment.

European Working Time Directive and doctors' health. A Systematic review of the available epidemiological evidence.

Running title: Long working hours and doctors' health

***Corresponding author**

Dr. Evangelia Demou
Healthy Working Lives Group
Institute of Health and Wellbeing
University of Glasgow
Glasgow, G12 8RZ
Tel: +44 141 330 3559
E-mail: evangelia.demou@glasgow.ac.uk

Maria Cruz Rodriguez-Jareño[1,2], Evangelia Demou[3,*], Sergio Vargas-Prada[4], Kaveh A Sanati[3,5], Alenka Škerjanc[6], Pedro G. Reis[7], Ritva Helimäki-Aro[8], Ewan B. Macdonald[3], Consol Serra[1,4,9,10], on behalf of the UEMS Section of Occupational Medicine[11].

1. Catalan and Spanish Societies of Occupational Medicine, Barcelona, Spain.
2. Department of Medical Sciences. School of Medicine. Universitat de Girona. Girona. Spain
3. Institute of Health and Wellbeing, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom
4. CiSAL - Centre for Research in Occupational Health, Universitat Pompeu Fabra, Barcelona, Spain.
5. Occupational Health Department, Epsom & St Helier University Hospitals NHS Trust, Surrey, United Kingdom
6. Clinical Institute of Occupational, Traffic and Sports Medicine, University Medical Centre, Ljubljana, Slovenia.
7. College of Occupational Medicine, Portuguese Medical Association, Lisbon, Portugal.
8. Helsinki City Occupational Health Centre, Helsinki, Finland.
9. Occupational Health Service, Parc de Salut MAR, Barcelona, Spain.
10. CIBER of Epidemiology and Public Health. Spain.
11. UEMS Section of Occupational Medicine. http://www.uems-occupationalmedicine.org

Word Count: 4,~~667~~794

ABSTRACT

Objective: To summarise the available scientific evidence on the health effects of exposure to working beyond the limit number of hours established by the European Working Time Directive (EWTD) on physicians.

Design: A systematic literature search was conducted in PubMed and Embase. Study selection, quality appraisal and data extraction were carried out by independent pairs of researchers using pre-established criteria.

Setting: Physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.

Participants: The total number of participants was 14,338.

Primary and secondary outcome measures: Health effects classified under the International Classification of Diseases (ICD-10).

Results: Over 3,000 citations and 110 full articles were reviewed. From these, eleven studies of high or intermediate quality carried out in North America, Europe and Japan met the inclusion criteria. Six studies included medical residents, junior doctors or house officers and the five others included medical specialists or consultants, medical, dental and general practitioners and hospital physicians. Evidence of an association was found between percutaneous injuries and road traffic accidents with extended hours/day or very long working weeks (LWH). The evidence was insufficient for mood disorders and general health. No studies on other health outcomes were identified.

Conclusions: LWH could increase the risk of percutaneous injuries and road traffic accidents, and possibly other incidents at work through the same pathway. While associations are clear, the existing evidence does not allow for an established causal or "dose-response" relationship between LWH and incidents at work, or for a threshold number of extended hours above which there is a significantly higher risk and the hours physicians could work and remain safe and healthy. Policy makers should consider safety issues when working on relaxing EWTD for doctors.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- The first systematic review, to our knowledge, on the effects of Long Working Hours (LWH) on physicians' health and safety.
- A systematic literature search conducted in PubMed and Embase with over 3,000 citations and 110 full articles reviewed.
- Eleven studies of high or intermediate quality carried out in North America, Europe and Japan, involving physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.
- The findings of this review confirmed that long working hours are associated with both an increased risk of needlestick injuries and road traffic accidents. Evidence was assessed as low or insufficient for mood disorders and general health.
- This review uncovered the lack of literature on the effect of long working hours on the general health of doctors, a topic that has important health ramifications not only for physicians but for patients too.

KEY WORDS: Working time, "Physicians"[MeSH], "Morbidity"[MeSH], "Occupational Injuries"[MeSH], Systematic review.

INTRODUCTION

The European Working Time Directive (EWTD),[1] applicable to all occupations across the EU, requires a maximum working week of 48 hours and establishes rest periods.[2, 3] Since it came into force in healthcare in 2009, it has been associated with concerns about the provision of health services including continuity of care, lower staffing levels, introduction of shift working, a reduction in training time and the adequate supervision of junior doctors.[4, 5] The large inter-country variations in physicians working hours as well as the mandatory controls over work hours across occupations raises further questions as to the basis of selecting and setting these limits and restrictions.[6] There is evidence of variable compliance to the EWTD within health care across Europe.[4, 5] The medical profession is also increasingly feminised with potential implications for women of childbearing age, which may coincide with the time in the profession where long working hours (LWH) are more likely to happen (i.e. during residency). Constraint on public-sector finances and increasing health care demands, have stimulated the debate on physicians' working hours,[7] with some evidence showing that the quality of health care is positively correlated with the number of physicians available to deliver that care.[8]

The Union of European Medical Specialists (UEMS) was created in 1958 to represent medical specialists in the EU, promote a high standard of training and practice at European level and facilitate the free movement of physicians across European countries.[9] The Occupational Medicine section was created in 1997 and is involved in issues related to specialist training, professional practice and research;[10] elaborating and publishing reference documents,[11] conducting and contributing to surveys,[12-14] and working in partnership with other key European groups.

The trigger for this project was a request by the Council of the UEMS to its Occupational Medicine Section for a position statement on the EWTD and possible health consequences to physicians of a potential removal of this professional group from the current EWTD 48 hours per week limit~~. An evidence based approach was taken by the Section and a systematic review of the literature was undertaken.~~[15-17] An evidence based approach was taken by the Section and a systematic review of the literature was undertaken.[15]

LWH have been associated with several adverse effects on workers' health, such as hypertension, cardiovascular disease, stress, depression, musculoskeletal disorders,

chronic infections, diabetes, general health complaints and all-cause mortality in a variety of occupational groups.~~[15-23]~~[18-26] In physicians, there is some evidence suggesting that LWH can result in impacts such as stress, depression, burnout, injuries, fatigue and sleep deprivation,~~[24-26]~~[27-29] and overwork has been considered one of the most stressful features of physicians' work.~~[24]~~[27] Some evidence exists of the beneficial effect of reducing the number of working hours on the quality of life of medical residents ~~[27]~~[30] and on burnout.[~~24~~27]

Field Code Changed
Field Code Changed

However, the relationship between work and health is complex. There is a substantial body of evidence showing that worklessness is associated with poorer health, whereas work is generally good for health and well-being,[31, 32] provided that the work environment is reasonably acceptable and supportive.[33]

Occupational and non-occupational exposures may play an important role such as the pattern and distribution of working hours, breaks, and recovery periods. Also, psychosocial and organisational factors, such as workload, job control, managers and peer support, training opportunities, individual characteristics and attitudes may be important.[6, 34]

Finally, physicians are a highly qualified, devoted and motivated professional group. Despite other existing reviews of a more general occupational approach, a review focused on physicians is warranted.

The aim of this project was to systematically review whether long-working hours (LWH) -i.e. defined as more than the 48 hour per week limit imposed by the EWTD- are associated with health effects –i.e. classified under the International Classification of Diseases (ICD-10) ~~[32]~~[35]- in physicians, and to examine what these associations are. This review does not include the potential impact of LWH on patient care or physician training.

METHODS

The systematic review sought to respond to the following research question: are LWH associated with health effects in physicians? For the purpose of this review, exposure to LWH was defined as working for more than 48 hours per week.~~[1]~~[1] Health outcomes included any disease as defined in the ICD-10 ~~[32] and work related injuries.~~[35] and work related injuries. If self-assessed tools on health status were used, only those validated were considered. Burnout, stress and outcomes such as individual symptoms, signs or biological markers (blood pressure, ECG, etc.) were not considered. Physicians of any medical, surgical or community speciality, working in any possible setting (hospitals, primary health care, etc.), as well as trainees, residents, junior house officers or postgraduate interns, were included.

Study identification

Electronic searches were carried out using PubMed and EMBASE as search engines (December 2011). Our search strategy was similar in both databases and consisted of a combination of the following keywords and Mesh terms: night shift, morning shift, evening shift, afternoon shift, shift work, rotating shift, shift combination, shift duration or length, shift system, clockwise rotation, shift roster, extended shifts, night work, evening work, work schedule, work hours, starting time, early start, irregular working hours, direction of rotation, overwork, extended hours, shift rota, workload, work schedule tolerance, sleep deprivation, sleep disorders, chronobiology disorders, circadian rhythm, psychomotor performance, circadian disruption, vigilance, alertness, wakefulness, drowsiness, fatigue, insomnia, hypersomnolence, dyssomnia, eveningness, morningness, neurocognitive performance, concentration difficulties, arousal, health, morbidity, mortality, disease, illness, stress, strain, distress, accident, injur*, death, suicid*, education, medical, physician, medical staff, hospital, doctor, surgeon, house officers, medical school, surgery, surgical.

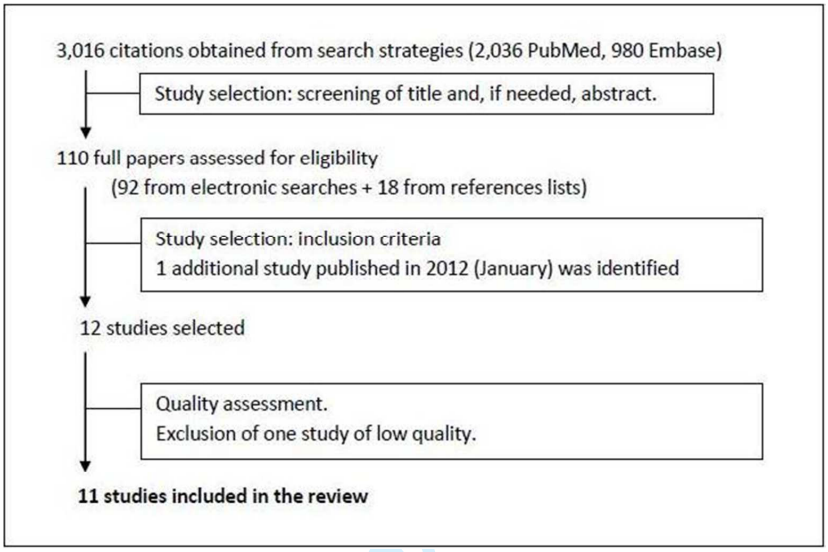
In addition, the references lists of articles selected for inclusion were carefully reviewed to identify additional studies.

Study selection

Studies were included if they measured the association between the exposure to LWH and health effects in physicians, used an observational epidemiological (i.e. cohort, case-control, cross-sectional) or experimental design~~, and were published in English, French, German, Italian, Slovenian or Spanish~~. Articles relating to on-call, night or shift work, but with unknown exposure or exposed to less than 48 working hours per week were excluded. They were also excluded if the working groups were other than physicians or addressed only other exposures or outcomes (i.e. patient safety, fatigue, sleep deprivation, social or family disruption). No limitation was set for languages.

A total of 2,036 citations were obtained from PubMed and 980 non-duplicated citations using Embase, yielding a total of 3,016 citations that were all screened by independent pairs of reviewers. All titles and, where necessary, abstracts were reviewed. Ninety two potentially suitable publications were identified from the electronic searches, and 18 further studies were identified from the references lists, yielding a total of 110 studies ~~for all of which~~published in English, French, German, Italian, Slovenian and Spanish. For all of them the full text was obtained and reviewed by independent pairs of reviewers. Disagreements within pairs were resolved by discussion and, where necessary, by a third reviewer who made the final decision. All authors participated as reviewers for screening the citations and full papers using well defined and pre-established criteria. Ninety nine articles were excluded at this stage. One further study was identified that was published after the search period. Finally, 12 papers were considered for quality assessment (Figure 1).

Figure 1. Results of the search strategy, using search engines on PubMed and EMBASE (December 2011) and screening of references lists of identified full papers, study selection and quality assessment.



The methodological quality of the 12 studies was assessed by independent pairs of reviewers. A standardized 16 point scale based on CONSORT and STROBE statements (Table 1) and adapted from a previous systematic review ~~[33] was used.~~ [36] was used. It includes 16 items grouped into 6 areas: 1) objectives, 2) study design, 3) target population and sample, 4) variables, 5) data sources, collection and measurement, and 6) statistical methods. Each item was rated as 1 (the requirement was met), 0.5 (the requirement was partially met) or 0 (the requirement was not or unclearly met). Disagreements within pairs of reviewers were resolved by consensus or, where necessary, by a third reviewer. For each of the 12 studies, a final score based on the sum of all items was assigned and the percentage was calculated based on a maximum score of 16. Study quality was rated as low, moderate or high if it scored less than 60%, between 60% and 79.9%, and 80% or more of the maximum score, respectively.

~~One article of low quality was excluded [34] and a total of 11 studies of intermediate or high methodological quality were included for the purpose of this review (figure 1).~~

One article of low quality was excluded [37] and a total of 11 studies of intermediate or high methodological quality were included for the purpose of this review (figure 1).

Data extraction

Selected information was extracted from each paper, including publication year, country of origin, study design, setting, study population, sample size, response rate, measure of working hours, definition used for LWH, health outcomes and their measurement, and

main results on point risk estimates or frequencies of health outcomes, and their corresponding 95% confidence interval (95%CI), and whether the analyses were adjusted for potential confounders.

Evidence synthesis

To summarise the results on the relationship between LWH and health effects, levels of evidence synthesis was performed. This was based on the methodological quality, study design and the consistency of the study outcomes. The following criteria were based on two previous relevant systematic reviews.[~~35~~38, ~~36~~39]

Field Code Changed

Field Code Changed

Strong evidence: consistent results in more than two studies of high quality.

Moderate evidence: consistent results in one high quality study and one intermediate, or between some studies of intermediate quality.

Insufficient evidence: identification of only one study or inconsistent results across studies;

Evidence of no association: consistent results of a non-association in two or more studies.

A measure of the possible magnitude of the association was attempted using the following criteria based on the association point estimate (RR = relative risk, rate ratio or odds ratio) when available or otherwise the comparison between two frequencies was used:

- high: RR > 3·00
- intermediate: RR = 1·50-3·00
- low: RR = 1·01-1·49
- no association: RR=1·00
- unclear

RESULTS

We identified 11 studies related to LWH and health effects in physicians that were eligible for inclusion. The outcome of the methodological quality assessment is given in table 1.

Methodological quality was appraised as high for 3 of the 11 studies, with scores ranging from 84 to 94%. The other 8 studies were considered as intermediate quality as

their scores ranged from 63 to 78%. The majority of all included studies received positive scores on items describing the study objectives, design and the study population (items from areas 1, 2 and 3), as well as the description of exposure and outcome variables (item 4). However, although data sources and collection were well described, reliability and validity of exposure to LWH and health outcome measures had low scores (items 5b and 5c). The statistical methods were in general appropriate, however the confounding variables were not measured in several studies and, as can be seen, the studies did not describe how they addressed missing data (items 6a and 6b).

Study characteristics:

The characteristics of the included studies are described in table 2.

Table 1. Methodological quality appraisal of the included studies

Study ID			Ayas 2006 [37][40]	Barger 2005 [38][41]	Firth-Cozens 1987 [39][42]	Fisman 2007 [40][43]	Hayasaka 2007 [41][44]	Kirkcaldy 1997 [42][45]	Rosta 2011 [43][46]	Stamp 2005 [44][47]	Sundquist 2000 [45][48]	Varma 2012 [46][49]	Zahrai 2011 [47][50]
1	Objectives	Are the objectives or hypotheses of the research described in the paper?	1	1	1	1	1	1	1	1	1	1	1
2	Study design	Is the study design presented?	1	1	1	1	1	1	1	1	1	1	1
3a	Target population	Do the authors describe the target population they wanted to research?	1	1	1	1	1	1	1	1	1	1	1
3b	Sample	Was a random sample of the target population taken? AND was the response rate 60% or more?	0.5	0.5	1	0.5	0.5	0.5	0.5	1	1	1	0.5
3c	Sample	Is participant selection described?	1	1	1	1	0.5	1	0	1	1	1	1
3d	Sample	Is participant recruitment described, or referred to?	1	1	1	1	1	1	1	1	1	1	0
3e	Sample	Are the inclusion and/or exclusion criteria stated?	1	1	1	1	0.5	0.5	1	1	1	1	0.5
3f	Sample	Is the study sample described? (minimum description: sample size, gender, age and occupation)	1	1	0	1	1	1	0	0	0	1	1
3g	Sample	Is the number of participants at each stage of the study reported?	1	1	1	0.5	1	0.5	1	1	1	1	1
4	Variables	Are the measures of long working hours and the health outcome described?	1	1	0.5	1	1	1	1	1	1	1	1
5a	Data sources, collection	Do authors describe the source of their data (e.g., official registry, health survey) AND how the data were collected?	1	1	1	1	1	1	1	1	1	1	1
5b	Measurement	Was reliability of the measure(s) of long working hours mentioned or referred to?	0	0	0	1	0	0	0	1	0	1	0
5c	Measurement	Was the validity of the measure(s) of long working hours mentioned or referred to?	1	1	0	0	0	0	0	0	0	0	0
5d	Measurement	Were health outcomes assessed by objective measures or validated self-reporting instruments?	0	1	1	1	1	0	1	1	1	1	1

6a	Statistical methods	Were appropriate statistical methods used and described, including those for addressing confounders?	0.5	1	0.5	1	0.5	0.5	1	0.5	1	1	0.5
6b	Statistical methods	Were the numbers/% of participants with missing data for long working hours and the health outcome indicated AND If more than 20% of data in the primary analyses were missing, were methods used to address missing data?	0	0	0	0.5	0	0	0	0	0	1	0.5
Total score** (%)			12.0 (75)	13.5 (84)	11.0 (69)	13.5 (84)	11.0 (69)	10.0 (63)	10.5 (66)	12.5 (78)	12.0 (75)	15.0 (94)	11.0 (69)
Quality rate***			interm	high	interm	high	interm	interm	interm	interm	interm	high	interm

** Maximum score = 16; *** Quality rate (%): low=<60; intermediate (interm)=60-79; high=80-100.

Table 2. Characteristics of the included studies.

Study ID	Country Of Study Population	Design*	Setting	Participants and sample size (response rate)	Working hours	Health outcomes	Quality score (%)
~~Ayas 2006~~ ~~[37]~~Ayas 2006 [40]	United States	~~Cross-over~~Cohort, prospective (1 year) / Case-crossover	Hospital	Interns in postgraduate residency programs Sample size: 2,737 (80%)	Mean (SD) hrs worked/month: 249.8 (75.3~~:self~~). Self-reported, monthly survey. Strong correlation (Pearson *r*=0.76; *p* 0.001) with hours worked 244 (69.3) from work diaries of randomly selected subset of 192 interns. Extended periods ~~(~~>(20 ~~hrs/day~~or more consecutive hours) vs. non extended periods (~~first~~ 12 hrs ~~at work~~or less consecutive hours).	Self-reported percutaneous injuries	12.0 (75)
~~Barger 2005~~ ~~[38]~~Barger 2005 [41]	United States	Cohort, prospective (1 year)	Hospital	Residents first postgraduate year (interns). Sample size: 2,~~554 (75~~737 (80%)	~~Average 70.7±26.0~~Mean (SD) hrs worked/week~~.~~ 70.7 (26.0). Extended shift (≥ 24hrs) vs. non extended shift (<24 hours). Self-reported, validated.	Self-reported and documented motor vehicle crashes.	13,5 (84)
~~Firth Cozens 1987~~ ~~[39]~~Firth Cozens 1987 [42]	United Kingdom	Cohort, prospective (1 year)	Hospital	Junior house officers Sample size: 170 (72%)	Mean number of hours per week = 90.6 hours (include on call). Self-reported.	GHQ-12 (case: score ≥2) and SCLDS, self-reported	11.0 (69)
~~Fisman 2007~~ ~~[40]~~Fisman 2007 [43]	United States and Canada	~~Cross-over~~Case-Crossover	Hospital	Medical trainees vs. other HCW. Sample size: 109 vs. 241 (46%)	Median number of hours per week: medical trainees = 70; other HCW = 40 (p<0.001). Self-reported, high reliability.	Reported percutaneous injuries to employee health care service.	13.5 (84)
~~Hayasaka 2007~~ ~~[41]~~Hayasaka 2007 [44]	Japan	Cross-sectional	Hospital, clinics, other	Female physicians Sample size: 367 (63%)	Comparison of increasing number of hrs/week from ≤ 30 to >50, self-reported.	GHQ-30 (case: score ≥8), self-reported	11.0 (69)
~~Kirkcaldy 1997~~ ~~[42]~~Kirkcaldy 1997 [45]	Germany	Cross-sectional	Not specified	Medical and dental practitioners. Sample size: 2,500 (not specified)	Long hours: mean 58.36 hrs/week SD 9.16 Short hours: mean 38.17 hrs/week SD 7.72 Self-reported.	Self-reported traffic accidents	10.0 (63)
~~Rosta 2011~~ ~~[43]~~Rosta 2011 [46]	Germany and Norway	Cross-sectional	Hospital	Hospital physicians. ~~Sample size~~Participants: 1,917 (58%) in Germany and 1,072 (65%) in Norway answered the questionnaire. From survey responders, 1,260 (65.7%) Germans and 562 (52.4) Norwegians ~~(58% and 65%,)~~were included in the analysis after applying further inclusion criteria.	German vs. Norwegian physicians (%): Hours per day >9 = 58.8 vs. 26.7 60 hours on-call per month = 63.4 vs. 18.3 Self-reported.	Validated questionnaire on self-rated health	10.5 (66)

Formatted: Font: (Intl) Arial Unicode MS
Formatted: Font: (Intl) Arial
Formatted: Font: Font color: Black, (Intl) Arial
Formatted: Font: Font color: Black, (Intl) Arial
Formatted: Font: Font color: Black, (Intl) Arial
Formatted: Font: Font color: Black, (Intl) Arial

~~Stamp 2005~~ [44]Stamp 2005 [47]	United States	Before-and-after	Hospital	Residents of general surgery Sample size: 28 (97%)	Changes of work patterns, after reduction from 90-110 to 78-80 work hrs/week.	SF-36, mental health; BDI, depression; self-reported	12.5 (78)
~~Sundquist 2000~~ [45]Sundquist 2000 [48]	Sweden	Cross-sectional	Primary care	General practitioners Sample size: 1,004 (72%)	Overtime defined as working at least 47 hrs/week. Self-reported.	Swedish SF-36, impaired mental health, self-reported	12.0 (75)
~~Varma 2012~~ [46]Varma 2012 [49]	Denmark	Cohort, prospective (20 months)	Hospital	Senior medical consultants Sample size: 2,790 (62%)	Long work hours (>40hrs/week). Self-reported.	Depression: redemption of anti-depressive drug prescriptions	15.0 (94)
~~Zahrai 2011~~ [47]Zahrai 2011 [50]	Canada	Cohort, prospective (6 months)	Hospital	Orthopaedic surgery residents Sample size: 16 (not specified)	Night float (n=9): 77.8% did >80h/week at baseline; 71.4% at follow-up). Standard call (n=7): 57.1% did >80h week at baseline; 80% at follow-up.	SF-36, mental health score, self-reported	11.0 (69)

* Follow-up period in brackets for prospective cohort studies; GHQ-12= General Health Questionnaire-12 items; SCLDS= Symptom Checklist Depression Scale; BDI = Beck Depression Inventory II.

Five had been carried out in North America (4 in the United States ~~[37, 38, 40, 44]~~[40, 41, 43, 47] and 1 in Canada [50]), five in European countries (Denmark,[49] Germany,[45] Norway and Germany [43]; Sweden,[48] and the United Kingdom [42]) and one study in Japan.[44] Four used a prospective cohort design (follow-up ranging from 6 to 20 months), 2 were case-crossover, 4 cross-sectional, and 1 was a before and after study. Overall, the total number of participants was 14,338 and included medical residents, junior doctors or house officers only working in hospitals, medical specialists or consultants, medical and dental practitioners, general practitioners, hospital physicians and one study included female physicians from a community service. Most studies (n=9) compared different working hour patterns within the same group of physicians, one compared physicians from two different countries,[46] and one included other health care workers as the comparison group.[43]

LWH was defined as more than 48hours/week in most studies, and some did not provide a definition. Two studies considered LWH below 48/week.[46, 48] All the included studies investigated health effects associated with working more than 48h/week, with number of hours ranging between <30 to 110hours/week. One paper studied the effects of an increasing number of hours from 30 to more than 50h/week.[44] Others studied more than 40h/week,[49] 47h/week,[48] above 58h/week [46] (two studies) and 6 studies referred to more than 70h/week,~~[37 40, 44, 47]~~[40-43, 47, 50] one of which compared 78-80h/week to 90-110 weekly hours.[47] Information on working hours was self-reported in nine studies, with reasonably good validity only in two studies ~~[37, 38]~~[40, 41] and reliability in one.[43] In the other two studies it was obtained from established work schedules.~~[44, 47]~~[47, 50]

Identified health outcomes that fulfilled the inclusion criteria were percutaneous injuries, motor vehicle accidents, mood disorders and general health. Information was self-reported in all studies using well known validated questionnaires, except three studies that used documented information on motor vehicle crashes,[40] reported percutaneous injuries [42] and antidepressant prescription data as a surrogate of depression.[49]

Summary of findings:

Table 3 shows the findings from each of the 11 studies.

Table 3. Results on the association between long working hours and identified health effects in physicians.

Health outcome	Main results	Adjustment by confounders	Study ID
Percutaneous injuries	Odds ratio (95% CI) for injuries during extended (>20 hours/day) vs. non extended periods (first 12 hours at work): All percutaneous injuries: 1.61 (1.46-1.78); Injuries reported to OH: 1.83 (1.48-2.28); Injuries in the ICU: 1.87 (0.69-5.04); Injuries in the operating room or labour and delivery: 1.77 (1.49-2.09); Injuries in the ICU, non-ICU, or ED: 2.17 (1.56-3.00)	Not adjustedTime of day and circadian influences	Ayas 2006[37]Ayas 2006[40]
Percutaneous injuries	Medical trainees vs. other health care workers: total median working hours: 70 vs. 40 hours (p<.001); median previous working hours at the time of injury: 6.5 vs. 5 hours (p< .001). Odds ratio (95% CI) for self-reported fatigue>: work > 40h in the past week: 3.550 (2.06-5.92); work during more than 5d in past week: 4.220 (4.46-7.15); >12h at work before injury: 8.58 (3.7-19.86). Incidence rate ratio (95%CI) for association between fatigue and injuries and fatigue: all workers: 1.40 (1.03-1.90); medical trainees: 2.94 (1.71-5.07); other HCWs: 0.97 (0.66-1.42) (p =0.001 for heterogeneity).	Age and sex	Fisman 2007[40]Fisman 2007[43]
Motor vehicle accidents	Relative riskOdds ratios (95% CI), after extended shift (≥ 24hrs) vs. non extended shift (<24 hours): Crashes = 2.3 (1.6-3.3); Near miss accidents = 5.9 (5.4–6.3)	Age and sex	Barger 2005[38]Barger 2005[41]
Motor vehicle accidents	Incidence rates of driving accidents on house visits: Long hours (mean 58.36 hrs/week) = 0.10 (0.31) vs. short hours (mean 38.17 hrs/week) = 0.02 (0.18) vs.). p<0.05	Not adjusted	Kirkcaldy 1997[42]Kirkcaldy 1997[45]
Mood disorders, depression	Hazard ratio (95% CI) hours/week intervals and redemption of anti-depressive drug prescription(reference group: 37-40 h/week): 25–36 h: 0.83 (0.24–2.82); 41–44 h: 0.95 (0.50–1.77); 45–49 h: 0.88 (0.43–1.78); 50–54 h: 0.83 (0.32–2.14); 55–59 h: 0.67 (0.15–2.94); >60 h: 0.48 (0.06–3.68). Cox regression analysis of work hours as a continuous variable: HR=0.93 (0.76–1.13)	Age, sex, marital status, medical specialty, decision authority at work, social support at work, quantitative work demands and previous redemption of AD drug prescription	Varma 2012[46]Varma 2012[49]
Mood disorders, GHQ-30	Prevalence of cases by working time (hours/week) ≤ 30 hours: 35.7% (41/115); >30 to 40 hours: 39.0% (39/100); >40 to 50 hours: 37.7% (26/69); >50 hours: 56.8% (46/81); *p*=0.0179 Stepwise multiple logistic regression, >50 h/wk. vs. ≤50 h/wk.: parameter 0.635 (p=0.0293)	Marital status, medical facility, position, and night duty.	Hayasaka 2007[41]Hayasaka 2007[44]
Mood disorders, GHQ-12 and SCLDS	No association was found between number of hours worked in a week and depression.	Not adjusted	Firth Cozens 1987[39]Firth Cozens 1987[42]
Mood disorders, SF-36	Scores when night float vs. standard call at baseline and follow-up: Mental health mean score (SD): baseline = 57.33 (22.63) vs. 65.71 (7.61); follow up = 52.00 (15.49) vs. 60.80 (11.45); p=0.72 Mental health component summary (SD), baseline = 34.84 (14.06) vs. 40.21 (7.61) ; follow up = 30.15 (10.71) vs. 42.40 (6.23); p=0.39 Regression analysis: increased number of hours in hospital correlated with significantly lower SF-36 scores in almost all domains.	Not adjusted	Zahrai 2011[47]Zahrai 2011[50]

Health outcome	Main results	Adjustment by confounders	Study ID
Mood disorders, SF-36 and BDI	Improvement of levels before and after duty hours limitations from 90-110 to 78-80 hors/week): Mental health SF-36): no statistically significant improvement Individual questions BDI before and after duty hours limitations, only energy level statistically significant.	Not adjusted	~~Stamp 2005~~[44]~~Stamp 2005~~[47]
Mood disorders, SF-36	Odds ratio (95% CI)for working 47 hours per week or more: Impaired mental health, men: 1.59 (0.95-2.66); women: 1.86 (1.03-3.37)	Age and amount of time in practice	~~Sundquist 2000~~[45]~~Sundquist 2000~~[48]
General health, SF-36	SF-36: Scores when night float vs. standard call at baseline and follow-up: General health mean score (SD): baseline = 62.11 (17.47) vs. 77.57 (24.25); follow up = 56.43 (24.89) vs. 84.20 (16.50); p=0.41 Physical health component summary (SD): baseline = 46.16 (13.15) vs. 52.01 (13.33); follow up = 39.32 (9.80) vs. 56.15 (2.18); p=0.015 Vitality mean score (SD): baseline = 51.67 (14.58) vs. 51.43 (15.74); follow up = 48.57 (14.92) vs. 51.00 (10.84); p=0.20 Regression analysis demonstrated that the increased number of hours spent in hospital correlated with significantly lower general health, physical function, mental health, role emotional, social function and mental component summary scale scores (all p < 0.05)	Not adjusted	~~Zahrai 2011~~[47]~~Zahrai 2011~~[50]
General health, SF-36	SF-36, BDI: Improvement of levels before and after duty hours limitations from 90-110 to 78-80 hours/week): Physical health: no statistically significant differences.	Not adjusted	~~Stamp 2005~~[44]~~Stamp 2005~~[47]
General health, SF-36	SF-36: Odds ratio (95% CI) adjusted for age and time in practice, for working 47 hours per week or more: Impaired general health: men = 1.66 (1.00-2.77); women = 1.59 (1.00-3.17)	Age and amount of time in practice	~~Sundquist 2000~~[45]~~Sundquist 2000~~[48]
General health~~, SF-36~~	Self-rated health: Odds ratio (95% CI) of good self-rated health (logistic regression): Norwegian work time pattern**= 1.35 (1.03-1.77); working in Norway = 4.17 (3.02 - 5.73).	Age, sex and country of work	~~Rosta 2011~~[43]~~Rosta 2011~~[46]

**Not working more than 9 hours a day and having more than 60 hours a month on-call); GHQ-12= General Health Questionnaire-12 items; SCLDS= Symptom Checklist Depression Scale; BDI = Beck Depression Inventory II.

Percutaneous injuries: One study of high quality and one of intermediate quality, with non-adjusted analyses, showed consistent results in medical trainees. Ayas et al. found an increased risk of percutaneous injuries associated with working more than 20 hours/day compared to non-extended hour periods, except in intensive care units;~~[37] Fisman et al.~~[40] Fisman et al. found an association between self-reported fatigue and the number of working hours at the time of injury and a 3-fold increased risk of sharp injuries in medical trainees compared to other health care workers (median of working hours/week before the injury: 70 vs. 40;p<.001).[~~40~~43]

Motor vehicle accidents: Two studies of high and intermediate quality showed that long working weeks were associated with an at least two-fold increased risk. Barger et al. found adjusted increased risks for car crashes (OR 2·3; 95%CI 1·6-3·3) and near miss accidents (OR 5·9; 95%CI 5·4-6·3) associated with working extended shifts.~~[38] Kirkcaldy et al.~~[41] Kirkcaldy et al. showed that non-adjusted incidence rates of traffic accidents on house visits was five times as much when physicians worked for a mean of 58 hours/week compared to 38 hours/week (p<0,05).[~~42~~45]

Mood disorders: Six studies of intermediate and high quality found contradictory results for mood disorders. Three of them, all of intermediate quality, provided evidence of an association between LWH and mental health problems. The prevalence of GHQ-30 cases was higher for female physicians working above 50 hours/week than for those working 30 or less hours/week (p<0.05),~~[41]~~[44] and an increased adjusted risk of impaired mental health was found for men and women general practitioners who worked for 47 hours per week or more.~~[45] Zahrai et al.~~[48] Zahrai et al. found that increased number of hours spent by residents in hospital correlated significantly with lower mental health and mental component summary scale scores (SF-36).~~[47]~~[50] Three other studies, one of which of high quality, did not find an association between depression or mood disorders and LWH. Varma et al. did not find higher adjusted risks of depression associated with increased number of hours in senior medical consultants, even when they worked for more than 60 hours per week.~~[46]~~[49] Depression or mood disorders measured with GHQ-12, SCLDS or BDI, all validated tools for mental health, were not associated with increased number of working hours ~~[39]~~[42] nor improved after decreasing the number of hours from 90-110 to 78-80 hours per week in residents.[~~44~~47]

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

General health: Four studies of intermediate quality analysed the association between LWH and general health. Zahrai et al. found that increased number of hours spent by residents in hospital correlated significantly with poorer general health, physical function and vitality using SF-36.~~[47]~~[50] A comparative study of physicians in two different countries found that Norwegians showed higher non-adjusted prevalence of self-reported good health compared to physicians in Germany who worked longer hours.~~[43]~~[46] An increased adjusted risk of impaired general health was found for men and women general practitioners who worked for more than 47 hours per week.~~[45]~~[48] However, following the implementation of reducing the number of hours from an average of 90-110hours/week to 78-80hours/week did not lead to an overall improvement of residents self-reported physical health.[~~44~~47]

Field Code Changed

Field Code Changed

Synthesis of the evidence:

The levels of evidence synthesis obtained from the analysis of the 11 studies included in this review are shown in table 4.

Table 4. Available scientific evidence on the health problems associated with long working hours in physicians: levels of evidence synthesis.

Health outcome	Degree of evidence*	Magnitude of the association**	Studies ID
Percutaneous injuries	++	++	~~Ayas 2006[37], Fisman 2007[40]~~Ayas 2006[40], Fisman 2007[43]
~~MMotor~~Motor vehicle accidents	++	++	~~Barger 2005[38], Kirkcaldy 1997[42]~~Barger 2005[41], Kirkcaldy 1997[45]
Mood disorders	+	+/-	~~Varma 2012[46], Hayasaka 2007[41], Firth Cozens 2001[39], Zahrai 2011[47], Stamp 2005[44], Sundquist 2000[45]~~Varma 2012[49], Hayasaka 2007[44], Firth Cozens 2001[42], Zahrai 2011[50], Stamp 2005[47], Sundquist 2000[48]
General health	+	+/-	~~Zahrai 2011[47], Stamp 2005[44], Sundquist 2000[45], Rosta 2011[43]~~Zahrai 2011[50], Stamp 2005[47], Sundquist 2000[48], Rosta 2011[46]

Formatted: Indent: Hanging: 0.1"

* Strong evidence (+++): consistent results in more than 2 studies of high quality; Moderate evidence (++): consistent results in two studies of high quality, or one high quality study and one intermediate, or between some studies of intermediate quality; Insufficient evidence (+): identification of only one

study or inconsistent results across studies; Evidence of no association (-): consistent results of a non-association in two or more studies.

** Magnitude of the association: it refers to the magnitude of the association point estimate (RR = relative risk, rate ratio or odds ratio): high (+++) if RR > 3; intermediate (++) if RR = 1·5-3; low (+) if RR < 1·5; no association (-); unclear (+/-).

There is moderate evidence of an association between LWH and percutaneous and motor vehicle accidents in physicians. This evidence comes from one study of high quality and another of intermediate quality both for percutaneous injuries, and for motor vehicle accidents and the magnitude of this association could be intermediate.

For mood disorders and general health there is a low or insufficient level of evidence of an association with LWH in physicians. This comes from 6 studies on mood disorders and 4 on general health of high and intermediate methodological quality and inconsistent results among them. No conclusion can be drawn about the magnitude of such associations if existed.

DISCUSSION

This review found moderate scientific evidence for a positive association of intermediate magnitude between exposure to LWH and percutaneous injuries and motor vehicle accidents in physicians. Evidence was assessed as low or insufficient for mood disorders and general health.

To our knowledge, this is the first systematic review on the effects of LWH on physicians' health and safety. The extensive searches were restricted to indexed journals. We systematically identified, selected and assessed the methodological quality of studies by means of independent pairs of reviewers. The quality assessment form and the levels of evidence used in this review were based on CONSORT and STROBE statements and on previous systematic reviews.~~[33, 35, 36]~~[36, 38, 39] To give some estimate of the potential magnitude of the effect, we added levels based on the point estimates of measures of association provided by the studies. However, such levels had not been previously established.

Strengths and limitations of the studies

Four studies had a longitudinal prospective design, however only two of them used a robust methodology. Two other studies used a crossover design, which reduces the likelihood of individual variability and confounding, and were considered well designed. Four studies, all rated as intermediate methodological quality, were cross sectional in design and therefore it is difficult to draw causal relationships from them. Although reverse causality cannot be ruled out from cross-sectional designs, it is unlikely that poorer health determines longer working hours than their healthier peers. One study used a quasi-experimental design but lacked a control group. Also, because of different designs used across the included studies, different estimates of the effects were used that may not be directly comparable.

Other limitations include the variability, validity and reliability of working hours, which in most studies was self-reported. Although only those papers specifying the number of working hours were accepted, shift and night work might have worked as confounders too, as it is difficult to acknowledge which proportion of the health effects observed was due exclusively to the LWH component and not due, for instance, to stress, fatigue or sleep deprivation. Moreover, the studies did not discriminate between time spent actively working or asleep while on call, though the EWTD considers all hours on call as working time. Likewise, it is possible that participants were aware that the study was conducted to assess associations between the number of hours worked and health outcomes. Therefore, there might be a volunteer bias in those doctors who work extended hours. To minimise this bias, a case-cross over analysis has been performed in three of the included study [40, 41, 43].

Formatted: Font: (Intl) Arial Unicode MS

Health outcomes were measured mainly based upon self-report, with the exception of documented motor vehicle crashes,~~[37]~~[40] reported percutaneous injuries ~~[39]~~[42] and register-notified anti-depressive drug prescriptions.~~[45] No studies with other objective measures of health (e.g.,~~[48] No studies with other objective measures of health (e.g., mortality) or mental disease (e.g., hospital data) were identified. However, self-reported or perceived health was assessed by validated and widely used instruments, especially when health status (general, mental or physical health) or ill-health symptoms are evaluated. Nonetheless, this raises the possibility that the observed associations might reflect differences in propensity to report health problems when they occur rather than true differences in the risk of worse health status. We cannot rule this out, but the higher propensity for reporting among those who work for longer hours and poorer work

schedule patterns would have to relate specifically to the reporting of ill-health rather than non-ill health. It seems unlikely that major differences in propensity to report would extend to a more concrete outcome.

The majority of the studies of this review took into account confounding variables in their analyses, such as sex and age. However, none of the studies analysed the potential effect of modifying factors, such as psychosocial aspects at work, including attitudes, motivation, job requirements, demands and content, organisational climate, social relationships at work, work satisfaction, supportive organisation, or the relative number of physicians and other staff available for patient care. Neither were important aspects outside of work, such as life events or lifestyles, taken into account. The study of Japanese female physicians found significant differences by marital status with higher GHQ scores for those who were married.[~~41~~44] There are significant differences in the number of physicians per capita across different countries and this together with the structure and organisation of health services are potential confounders too that were not addressed in the selected studies across countries.

Field Code Changed

Field Code Changed

A relevant outcome from this review would be to establish a threshold number of extended hours above which there is a significantly higher risk. There is a lack of evidence for a dose-response relationship and nor does the evidence give any indication for a threshold number of hours that physicians could work and remain safe and healthy.

Finally, the heterogeneity of the included studies did not allow a more quantitative synthesis, and a level of evidence approach was used instead.

Comparison with other studies

Our findings of an association between LWH and injuries and accidents in physicians are consistent with research in other occupational groups. Dembe et al. carried out a longitudinal survey in the US including more than 10,000 workers from a variety of occupations and settings.~~[16]~~[19] A strong dose-response effect (adjusted for age, gender, occupation, industry and region) was found between LWH and injuries above 40h/week and 8h/day; jobs with overtime schedules were associated with a 61% higher injury hazard rate compared to jobs without overtime; working at least 12 hours per day was associated with a 37% increased hazard rate, and working at least 60 hours per week was associated with a 23% increased hazard rate. They concluded that LWH

might indirectly precipitate workplace injuries through a causal process by inducing fatigue or stress in workers. In a separate study assessing injury risks to health care personnel, Dembe et al. demonstrated that the risk of injury when working overtime or at least 60 h per week among physicians and nurses was statistically significant.[51] In our review we found that information about the length of shift work varied across studies, some analysing shifts of up to twenty four hours. Work that included such shifts was more strongly associated with accidents and injuries.[~~37~~40, ~~38~~41]

Field Code Changed

We found insufficient evidence of an association between LWH and mood disorders, other diseases and general health in physicians, despite some scientific evidence suggesting that LWH increase morbidity and mortality in other occupations.[~~16 23~~19-26]

Field Code Changed

Possible explanations for these inconsistencies might be the scarcity of high quality research, and the fact that some of the included papers did not fully address the effect of confounding factors, and the possibility that working as a physician might have a protective effect by itself, for instance through higher motivation and social recognition relative to other occupations.

Stamp et al., in the only before and after study, did not find significant differences for mood disorder or general health before and after implementing requirements to reduce working schedules to 80h/week in residents.~~[44]~~[47] It is debatable, however, whether any conclusions can be drawn from the impact of reducing extremely high working hours (90-110hours/week) to very high working hours (78-80hours/week).

~~Although we searched for and identified literature describing associations between LWH and increased mortality in doctors,[48, 49] none of those studies specified the number of hours worked, and therefore were not included.~~

Although we searched for and identified literature describing associations between LWH and increased mortality in doctors,[52, 53] none of those studies specified the number of hours worked, and therefore were not included.

Recommendations for future research

Further research, including well designed prospective and retrospective longitudinal studies, is needed in this field to establish on scientific grounds what number of hours

are safe, both to patients and physicians. Also studies on long term health effects should be considered.

This review was aimed to study exclusively possible detrimental health effects of LWH (more than 48 hours per week) to physicians. Other systematic reviews should be conducted addressing issues, such as the impact of working hours on doctor's well-being and quality of life, the quality of junior doctors training, the quality and continuity of care and impact on patient safety. Also, from a staffing perspective, the economic and social burden of sick leave and physicians leaving the profession are important issues that were beyond the scope of this review and that should be considered in future research. ~~Other established criteria [50]~~Other established criteria [54] that working hours should be 'family friendly', promote gender equality, enhance productivity or facilitate worker choice and influence over working hours, were not addressed. Further research is needed in these other areas.

Conclusions

The findings of this review have confirmed that long working hours are associated with both an increased risk of needlestick injuries and road traffic accidents. The fact that these associations are labelled moderate in this review, should not preclude safety procedures and measures being assessed and reassessed to ensure maximum physician and patient safety. The reduction of needlestick injuries requires improved use of safe sharp devices, compliance with safe working procedures, training and regular audits of working practices, particularly for junior doctors who are at greatest risk of needlestick injuries, possibly due to inadequate training.~~[51, 52]~~[55, 56] The risk of road traffic accidents could be reduced by discouraging driving after very long shifts (more than 16 hours) and where there is sleep deprivation.~~[53]~~[57] Transport should be provided by the employer in these situations. This is a pragmatic recommendation which recognises that long working hours have not been eliminated as a result of the EWTD.

This review uncovered the lack of literature on the effect of long working hours on the general health of doctors, a topic that has important health ramifications not only for physicians but for patients too. However, the paucity of available studies, which does not allow for a causal or dose-response relationship to be established, raises further questions about the evidence-base of the current 48 hour limit, especially as this is not

enforced strictly and junior doctors, for example in the UK, may opt-out.[54][58] It may be relevant that the EWTD was not based entirely on science but also took into account, as does all EU legislation, the views of "social partners".[3][3]

It is likely that the mechanism for these increased health and safety risks is fatigue, and our finding of increased risk of accidents to staff, may also be associated with increased risk of clinical errors. Patient safety was not the purpose of this review but long working hours are well recognised to cause decrement in performance both in health care and other professions, where the performance of staff can be safety critical, such as airline pilots, and professional drivers.[5559]

Field Code Changed

Field Code Changed

In addition the EWTD has been associated with real concerns about the adequacy of the training of doctors given their reduced exposure to patient care, and the lack of experience of the patient journey, imposed by adherence to the EWTD.[4][4] This is potentially exacerbated at the same time by the constraints on the overall duration of training imposed by other EU directives.[56][60] Further reviews of the EWTD, with regard to this unique group of workers, need to take into account other social factors such as the impact on patient care. This is particularly the case when all health care systems are under strain because of burgeoning demands and limited resources.[4][4]

The findings of this systematic review lead to the suggestion of the following guidelines in the event of non-compliance with the EWTD: 1) long shifts should be avoided to protect both physicians' health and patient safety and no shift should be longer than 16 hours (and then exceptionally);[28] 2) Physicians should be discouraged from driving after long shifts to reduce the risk of motor accidents;[41, 45] 3) rigorous attention should be paid to reducing the risk of sharps injuries;[40, 43] 4) organisational aspects such as workload and job control, as well as the pattern and distribution of working hours, breaks and recovery periods should be carefully taken into account, in order to avoid fatigue and sleep deprivation that could lead to mistakes and accidents;[61, 62] 5) Physicians should work in supportive psychosocial environments, in teams, and with adequate training and supervision as their wellbeing is important also for the health of their patients.[61, 62]

Further research is required to establish any longer term effects such as on mortality and mental health, also the impact of the nature of the work organisation taking into account

the psychosocial aspects of the physicians working and non-working lives, and to determine how many hours are safe for physicians to work.

ACKNOWLEDGMENTS

Leena Isotalo, Trial Search Coordinator at the Cochrane Occupational Safety and Health Review Group, designed and conducted the final search of the evidence using Medline and Embase databases.

CONFLICT OF INTERESTS

None. No financial relationships with any organisations that might have an interest in the submitted work, and no other relationships or activities that could appear to have influenced the submitted work.

FUNDING

The Section of Occupational Medicine of the UEMS provided limited support to the work of all authors (MCR, ED, SVP, KS, AS, PR, RH), with the exception of the senior authors (EBM, CS). The expenses of one meeting of the core group (MCR, ED, SVP, EBM, CS) were also funded by the Section. **Role of funding source:** The Section of Occupational Medicine of the UEMS acted in its role of promoting and contributing to the research in Occupational Medicine.

AUTHORS' CONTRIBUTIONS:

Mari Cruz Rodríguez-Jareño, Evangelia Demou, Sergio Vargas-Prada, Alenka Škerjanc, Kaveh A Sanati, Ewan B. Macdonald and Consol Serra conceived and submitted the design, carried out the acquisition of data, analysis and interpretation of data, drafted the article and revised it.

Pedro G. Reis and Ritva Helimäki-Aro contributed to acquisition of data, analysis and interpretation of data, drafted the article and revised it.

All authors gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

ETHICS COMMITTEE APPROVAL was not sought as no patients were involved in this study.

REFERENCES

1 EUR-Lex. Directive 2003/88/EC of the European Parliament and of the Council of 4 November 2003 concerning certain aspects of the organisation of working time. In: Union OJotE, ed. Brussels: Official Journal of the European Union, 2003.

2 EPHA. 2012 EPHA Briefing on the European Working Time Directive (EWTD) http://www.epha.org/IMG/pdf/EPHA_Briefing_EWTD_Oct2012_FINAL.pdf; Access Year 2012.

3 European Commission. 2013 Working Conditions-Working Time Directive. http://ec.europa.eu/social/main.jsp?catId=706&langId=en&intPageId=205; Access Date December 2013.

4 House J. Calling time on doctors' working hours. Lancet 2009;**373**:2011-12.

5 The Lancet. Doctors' training and the European Working Time Directive. Lancet 2010;**375**:2121.

6 Leff D, Aziz O, Darzi A. Trucks, planes, and scalpels - Is there an evidence-based approach to surgeons' working hours? Arch Surg 2007;**142**:817-20.

7 Parthasarathy S. Sleep and the medical profession. Curr Opin Pulm Med 2005;**11**:507-12.

8 Dr. Foster Hospital Guide. 2011 www.drfosterhealth.co.uk; Access Date February 2011.

9 UEMS. 2012 The Union of European Medical Specialists. www.uems.net; Acess Year 2012.

10 Macdonald EB, Ritchie KA, Murray KJ, Gilmour WH. Requirements for occupational medicine training in Europe: a Delphi study. Occup Environ Med 2000;**57**:98-105.

11 Macdonald E, Baransky B, Wilford J. Occupational Medicine in Europe: Scope and Cometencies. In: Health WECfEa, ed. Health, Environment and Safety in Enterprises Series n. 3. Bilthoven, The Netherlands, 2000.

12 Ballester M, Cordell N, Rodriguez Jareno MC, Serra C, Med USO. A European survey of professional bodies representing occupational medicine specialists. Occup Med (Oxf) 2012;**62**:366-70.

13 Cashman C, Slovak A. The occupational medicine agenda: Routes and standards of specialization in occupational medicine in europe. Occup Med (Oxf) 2005;**55**:308-11.

14 De Schryver A, Claesen B, Meheus A, van Sprundel M, Francois G. European survey of hepatitis B vaccination policies for healthcare workers. Eur J Public Health 2011;**21**:338-43.

~~15~~15 Rodriguez-Juareno M, Serra C, Demou E, et al. European Working Time Directive: systematic review for evidence based decision making. SUMMARY REPORT. In: Medicine USoO, ed. http://www.uems-occupationalmedicine.org/sites/default/files/old_userfiles/File%203%20-%20EWTD%20UEMS%20REPORT%2005%2010%202011.pdf,: UEMS Section of Occupational Medicine, 2011.

16 UEMS. 2010 The Newsletter of European Medical Specialists. http://www.uems.eu/__data/assets/pdf_file/0014/1463/1466.pdf.

17 UEMS. Meetings of the Board & Council of UEMS- Report. In: European Union of Medical Specialists; http://www.uems.eu/__data/assets/pdf_file/0011/1235/UEMS_2011.37_-_report_-_UEMS_Council_6-8.10.2010_Napoli.pdf, ed. Napoli (Italy), 2011.

18 Bannai A, Tamakoshi A. The association between long working hours and health: A systematic review of epidemiological evidence. Scand J Work Environ Health 2013.

~~16~~19 Dembe AE, Erickson JB, Delbos RG, Banks SM. The impact of overtime and long work hours on occupational injuries and illnesses: new evidence from the United States. Occup Environ Med 2005;**62**:588-97.

~~17~~20 Johnson JV, Lipscomb J. Long working hours, occupational health and the changing nature of work organization. Am J Ind Med 2006;**49**:921-9.

~~18~~21 Kivimaeki M, Batty GD, Hamer M, et al. Using Additional Information on Working Hours to Predict Coronary Heart Disease. Ann Intern Med 2011;**154**.

~~19~~22 Liu Y, Tanaka H, Fukuoka Heart Study G. Overtime work, insufficient sleep, and risk of non-fatal acute myocardial infarction in Japanese men. Occup Environ Med 2002;**59**:447-51.

~~20~~23 Sokejima S, Kagamimori S. Working hours as a risk factor for acute myocardial infarction in Japan: case-control study. BMJ 1998;**317**:775-80.

~~21~~24 Spurgeon A, Harrington JM, Cooper CL. Health and safety problems associated with long working hours: A review of the current position. Occup Environ Med 1997;**54**:367-75.

~~22~~25 Virtanen M, Ferrie JE, Singh-Manoux A, et al. Long working hours and symptoms of anxiety and depression: a 5-year follow-up of the Whitehall II study. Psychol Med 2011;**41**:2485-94.

~~23~~26 Virtanen M, Stansfeld SA, Fuhrer R, Ferrie JE, Kivimaki M. Overtime work as a predictor of major depressive episode: a 5-year follow-up of the Whitehall II study. PloS One 2012;**7**:e30719.

2427 Gopal R, Glasheen JJ, Miyoshi TJ, Prochazka AV. Burnout and internal medicine resident work-hour restrictions. Arch Intern Med 2005;**165**:2595-600.

2528 Reed DA, Fletcher KE, Arora VM. Systematic Review: Association of Shift Length, Protected Sleep Time, and Night Float With Patient Care, Residents' Health, and Education. Ann Intern Med 2010;**153**:829-42.

2629 Sargent MC, Sotile W, Sotile MO, Rubash H, Barrack RL. Quality of Life During Orthopedic Training and Academic Practice Part 1: Orthopedic Surgery Residents and Faculty. J Bone Joint Surg Am 2009;**91A**:2395-405.

2730 Fletcher KE, Underwood W, Davis SQ, Mangrulkar RS, McMahon LF, Saint S. Effects of work hour reduction on residents' lives - A systematic review. JAMA 2005;**294**:1088-100.

2831 Black C, Frost D. *Health at work - an independent review on sickness absence*. London: The Stationary Office, 2011.

2932 Waddell G, Burton K. *Is work good for your health and wellbeing?* London: The Stationary Office, 2006.

3033 Butterworth P, Leach LS, Strazdins L, Olesen SC, Rodgers B, Broom DH. The psychosocial quality of work determines whether employment has benefits for mental health: results from a longitudinal national household panel survey. Occup Environ Med 2011;**68**:806-12.

3134 Lockley SW, Landrigan CP, Barger LK, Czeisler CA. When policy meets physiology - The challenge of reducing resident work hours. Clin Orthop 2006:116-27.

3235 ICD-10. 2012 International Classification of Diseases. http://www.who.int/classifications/icd/en/; Access Year 2012.

3336 van Uffelen JGZ, Wong J, Chau JY, et al. Occupational Sitting and Health Risks A Systematic Review. Am J Prev Med 2010;**39**:379-88.

3437 Kirsling RA, Kochar MS, Chan CH. An evaluation of mood states among 1st-year residents. Psychol Rep 1989;**65**:355-66.

3538 Bernard B. A critical review of epidemiologic evidence for work-related musculoskeletal disorders of the neck, upper extremity and low back. In: NIOSH, ed. Publication No. 9741. Cincinnati: National Institute for Occupational Safety and Health, 1997.

3639 Steenstra IA, Verbeek JH, Heymans MW, Bongers PM. Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: a systematic review of the literature. Occup Environ Med 2005;**62**:851-60.

3740 Ayas NT, Barger LK, Cade BE, et al. Extended work duration and the risk of self-reported percutaneous injuries in interns. JAMA 2006;**296**:1055-62.

~~38~~41 Barger LK, Cade BE, Ayas NT, et al. Extended work shifts and the risk of motor vehicle crashes among interns. N Engl J Med 2005;**352**:125-34.

~~39~~42 Firth-Cozens J. Emotional Distress in Junior House Officers. BMJ 1987;**295**:533-36.

~~40~~43 Fisman DN, Harris AD, Rubin M, Sorock GS, Mittleman MA. Fatigue increases the risk of injury from sharp devices in medical trainees: Results from a case-crossover study. Infect Control HospEpidemiology 2007;**28**:10-17.

~~41~~44 Hayasaka Y, Nakamura K, Yamamoto M, Sasaki S. Work environment and mental health status assessed by the general health questionnaire in female Japanese doctors. Ind Health 2007;**45**:781-86.

~~42~~45 Kirkcaldy BD, Trimpop R, Cooper CL. Working hours, job stress, work satisfaction, and accident rates among medical practitioners and allied personnel. Int J Stress Manag 1997;**4**:79-87.

~~43~~46 Rosta J, Aasland OG. Work Hours and Self rated Health of Hospital Doctors in Norway and Germany. A comparative study on national samples. BMC Health Serv Res 2011;**11**.

~~44~~47 Stamp T, Termuhlen P, Miller S, et al. Before and after resident work hour limitations: an objective assessment of the well-being of surgical residents. Curr Surg 2005;**62**:117-21.

~~45~~48 Sundquist J, Johansson SE. High demand, low control, and impaired general health: working conditions in a sample of Swedish general practitioners. Scand J Pub Health 2000;**28**:123-31.

~~46~~49 Varma A, Marott J, Stoltenberg C, Wieclaw J, Kolstad H, Bonde J. With long hours of work, might depression then lurk? A nationwide prospective follow-up study among Danish senior medical consultants. Scand J Work Environ Health 2012;**38**:418-26.

~~47~~50 Zahrai A, Chahal J, Stojimirovic D, Schemitsch EH, Yee A, Kraemer W. Quality of life and educational benefit among orthopedic surgery residents: a prospective, multicentre comparison of the night float and the standard call systems. Can J Surg 2011;**54**:25-32.

~~48~~51 Dembe AE, Delbos R, Erickson JB. Estimates of injury risks for healthcare personnel working night shifts and long hours. Quality & Safety in Health Care 2009;**18**:336-40.

52 Lindfors PM, Nurmi KE, Meretoja OA, et al. On-call stress among Finnish anaesthetists. Anaesthesia 2006;**61**:859-64.

~~49~~53 Šelb J, Albreht T. Mortality rates of medical doctors in Slovenia in 1985 to 1999. Zdravniški vestnik 2000:I47-I48.

~~50~~54 Lee S, McCann D, Messenger J. Working Time Around the World: Trends in working hours, laws and policies in a global comparative perspective. Geneva, Switzerland: International Labour Office, 2007.

~~51~~55 Elder A, Paterson C. Sharps injuries in UK health care: a review of injury rates, viral transmission and potential efficacy of safety devices. Occup Med-Oxford 2006;**56**:566-74.

~~52~~56 Naghavi SHR, Sanati KA. Needlestick injuries: Does left-handedness matter? Am J Infect Control 2009;**37**:341-41.

~~53~~57 Connor J, Norton R, Ameratunga S, et al. Driver sleepiness and risk of serious injury to car occupants: population based case control study. BMJ 2002;**324**:1125-28A.

~~54~~58 BMA. 2013 European Working Time Directive, http://bma.org.uk/practical-support-at-work/ewtd,. Access Date December 2013.

~~55~~59 Wilson AM, Weston G. Application of Airline Pilots Hours to Junior Doctors. BMJ 1989;**299**:779-81.

~~56~~60 The European Specialist Medical Qualifications Order 1995. 1995 Statutory Instruments, 1995 No. 3208, MEDICAL PROFESSION, The European Specialist Medical Qualifications Order 1995, http://legislation.data.gov.uk/uksi/1995/3208/made/data.htm?wrap=true,. Access Date December 2013.

~~57~~61 Boorman S. NHS Health and Well-being. The Boorman Review. Leeds, 2009.

~~58~~62 Williams S, Michie S, Pattani S. Improving the health of the NHS workforce. Report of the partnership on the health of the NHS Workforce. London: The Nuffield Trust, 1998.

Formatted: EndNote Bibliography, Indent: Left: 0", Hanging: 0.19", Space Before: 0 pt, After: 0 pt, Line spacing: single

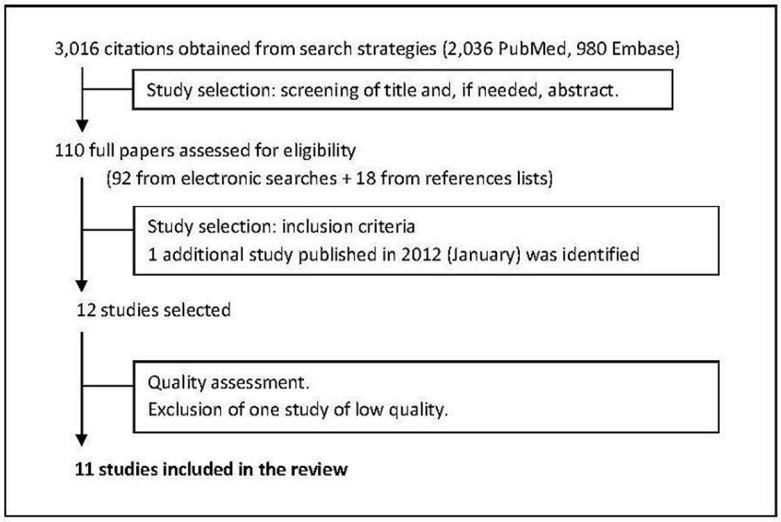


Figure 1.Results of the search strategy, using search engines on PubMed and EMBASE (December 2011) and screening of references lists of identified full papers, study selection and quality assessment.
90x59mm (300 x 300 DPI)

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6-7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	6
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	8-9
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	8-9
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	8-9
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8-9
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	9
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	8-9

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	7
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	10-11
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	12-13
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	18
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	9-19
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	19
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	19-21
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	21-23
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	24-25

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.