Systematic review

Systematic review and meta-analysis of the effect of North American working hours restrictions on mortality and morbidity in surgical patients

M. H. Jamal[1,2], S. A. R. Doi[3,4], M. Rousseau[1], M. Edwards[1], C. Rao[3], J. J. Barendregt[3], L. Snell[1] and S. Meterissian[1]

[1]Department of General Surgery and Center for Medical Education, McGill University, Montreal, Quebec, Canada, [2]Department of General Surgery, Kuwait Medical School, Kuwait City, Kuwait, [3]School of Population Health, University of Queensland and [4]Department of Endocrinology, Princess Alexandra Hospital, Brisbane, Queensland, Australia

Correspondence to: Dr M. H. Jamal, Department of Hepatobiliary and Multiorgan Transplantation, Royal Victoria Hospital, Montreal, Quebec, Canada H3A 1A1 (e-mail: u22mj@yahoo.com)

Background: **Short duty hours, imposed by the Accreditation Council of Graduate Medical Education (ACGME) regulations, have been claimed to be associated with loss of continuity of care among surgical patients, leading to a potentially increased risk of adverse surgical outcomes. This systematic review and meta-analysis assessed the strength of associations between duty hour restrictions and morbidity and mortality of various surgical procedures.**

Methods: **MEDLINE, Embase, BIOSIS Previews®, the Education Resources Information Center and the Cochrane Central Register of Controlled Trials (January 2000 to September 2009) were searched, and reports screened to identify comparative studies of mortality and morbidity before and after the introduction of ACGME regulation periods. Random-effects (RE) and quality-effects (QE) meta-analyses were performed to determine the risk of morbidity or death associated with long duty hours compared with shorter duty hours. Results are presented as odds ratio (OR) with 95 per cent confidence interval.**

Results: **A total of 19 data sets (10 articles), including 730 648 subjects in the mortality studies and 64 346 in the morbidity studies, were analysed. Long duty hours were associated with a non-significantly increased risk of death compared with shorter duty hours (OR 1·28, 0·94 to 1·73). There was no difference in morbidity between the two groups (OR 1·03, 0·67 to 1·57). Mortality associations were generally stronger for general surgery, more recent studies and higher-quality studies. Heterogeneity was evident among the studies included.**

Conclusion: **The reduction in working hours has not affected patient care negatively in terms of demonstrable differences in morbidity and mortality. However, it cannot be distinguished whether this effect is actually due to a non-detrimental effect of the reduction in working hours or whether any such detriment is offset by continually improving patient care and increased surgical supervision.**

Paper accepted 21 November 2011

Published online 12 January 2012 in Wiley Online Library (www.bjs.co.uk). **DOI:** 10.1002/bjs.8657

Introduction

The USA saw the first reduction in duty hours in the State of New York in 1989[1], followed by the Accreditation Council of Graduate Medical Education (ACGME) working hours restriction in 2003. The subject of duty hours restriction is still a very active area of research and debate, especially with the release of the recent ACGME task force report[2], suggesting a further reduction in working hours in the USA starting in 2011. This is also a global issue; on 9 June 2009, a review of the effect of the European Working Time Directive on the quality of medical training in England was released. The directive, which came into force in Europe on 1 August 2009, affects doctors in training by limiting the average working week to 48 h[3]. The Royal College of Surgeons has been critical of the directive, responding that 'the remedies proposed are unworkable'[3]. Indeed, surgeons in Europe have mixed perceptions from perceiving it as harming training[4] to more optimistic views suggesting that exposure to operative surgical training can be sustained despite shortening of working hours[5]. Overall, however, regulations concerning restricted working hours for surgeons have largely been criticized or disregarded[3,6,7].

Surgical specialties are the most affected by the duty hours restrictions owing to the large amount of emergency and out-of-hours work[2,3]. The long-held traditions of surgical patient care ensure maximum continuity, minimizing as much as possible the numbers of handovers and cross-coverage. Traditionally, surgical education relies on the continuous follow-up of patients from admission to discharge with participation in the operative intervention. Reducing duty hours potentially threatens these traditions, and has not been well received in the surgical community[8]. The fear has been that with reduction of continuity of care there will be a reduction in quality of care for surgical patients, resulting in worse outcomes as well as a decrease in quality of surgical training. Even though there is ample literature linking fatigue to medical errors and reduced outcomes[9], there is no study linking worse surgical patient outcome to surgical resident fatigue. With regard to surgical resident education and practice, a systematic review found maintenance of resident operating room experience as well as examination scores after introduction of the ACGME regulations. Improvement in resident quality of life and reduced stress levels were also demonstrated[1]. However, a reduced quality of life of surgical faculty members, with dissatisfaction with resident training and patient care, was found[1].

Eight years after institution of the ACGME duty hours regulations, there is a need to evaluate their effect on surgical patient outcome. The aim of this systematic review and meta-analysis was to investigate the effect of the introduction of working hours restrictions in North America on morbidity and mortality in surgical patients.

Methods

Data sources

The English and French medical and surgical literature, from January 2000 to September 2009, was searched for articles examining the impact of surgical residents' working hours on their personal well-being, education and training, on patient care and on faculty members. Databases included in the search were MEDLINE, Embase, BIOSIS Previews®, the Education Resources Information Center (ERIC) and the Cochrane Central Register of Controlled Trials. The search strategy is described in detail in *Appendix S1* (supporting information) and outlined in *Fig. 1*.

Eligibility criteria

Studies were included if they looked at the effect of the ACGME working hours restriction by comparing mortality and morbidity before and after implementation of the regulations, were conducted in a surgical residency programme in the USA, and included outcomes directly related to patient care.

Surgical specialties were defined as any specialty with operating room privileges. This included general surgery with its subspecialties, cardiac surgery, ear nose and throat, plastic surgery, ophthalmology, orthopaedics, obstetrics and gynaecology, and neurosurgery.

All opinion papers, editorials and irrelevant abstracts were excluded. Studies of European working hours were

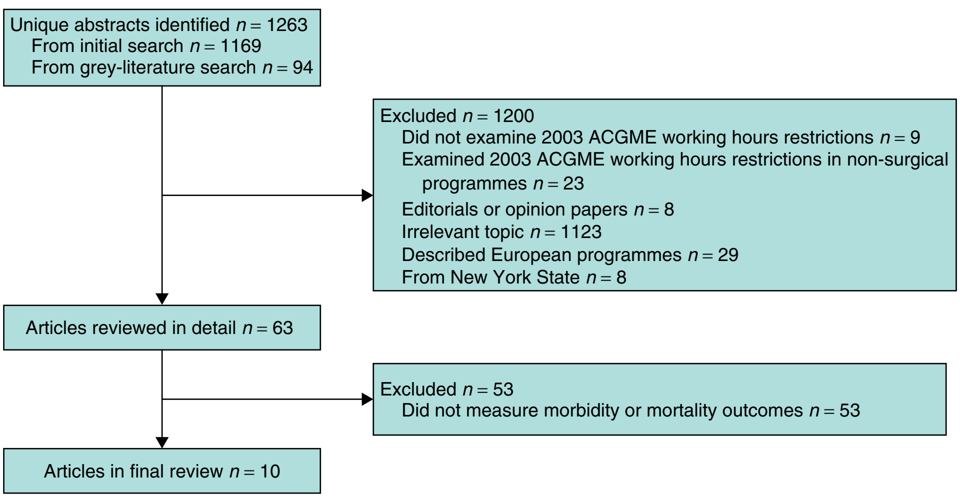

Fig. 1 Selection of studies for review. ACGME, Accreditation Council of Graduate Medical Education

excluded, as their limitations were very different from the ACGME limitations. All papers describing programmes in New York State were also excluded because duty hour restrictions started here in 1989 and were not monitored as effectively as the ACGME working hours.

Data abstraction and quality assessment

A standard data extraction form summarizing the study design, study population and relevant raw data was completed for each article by two independent reviewers. A quality scoring instrument was devised based on a system for grading evidence in medical education developed by Harden and colleagues[10] (*Table S1*, supporting information). Full details on the score creation process have been published previously[1].

Statistical analysis

Morbidity and mortality (binary) were determined by the various criteria used in individual studies. Although these studies were classified as retrospective cohort studies and the direction was forward, sample selection in all studies was random (not by exposure or outcome). Therefore, the mortality or morbidity odds ratio (OR) was used as a measure of the relationship between ACGME status and risk of the defined event. A value of 1 indicates equivalence, and a value greater than 1 indicates a higher risk of morbidity or mortality for the pre-ACGME groups.

Heterogeneity was determined to be present when the value of τ^2 was greater than zero[11] and/or the *Q*-statistic was significant at $P < 0{\cdot}1$. Heterogeneous studies were combined in both quality-effects (QE) and random-effects (RE) models. The QE model was also used to weight more commonly done procedures based on national estimates of principal procedure-only discharges in each group taken from the Healthcare Cost and Utilization Project Nationwide Inpatient Sample, 2008, Agency for Healthcare Research and Quality (AHRQ). This is based on data collected by individual states and provided to AHRQ by the states. For trauma, a comparable number used was that based on number of discharges for multiple specific diagnoses in the diagnosis-related group related to trauma. These were converted to population impact probabilities by dividing each population estimate by the largest population estimate among the studies included in the meta-analysis[12]. The final study-specific score for use in the QE model was this population impact probability multiplied by the probability of credibility of each study (quality score divided by its maximum value) to create a combined score (*Appendix S2*, supporting information[13–17]).

To test for publication bias, funnel plot asymmetry was checked. The funnel plot was considered to be asymmetrical if the intercept of Egger's regression line deviated from zero with $P < 0{\cdot}100$. Finally, to assess the robustness of the meta-analysis, sensitivity analyses were performed by examining the effect of subgrouping of the various studies on the pooled results. A one-way sensitivity analysis was also carried out by excluding individual studies from the meta-analysis. All analyses were done using MetaXL (http://www.epigear.com).

Table 1 Calculation of quality scores (incorporating population weights) for use in the meta-analysis

Procedure type	Points assigned based on quality checklist (maximum 17)	Probability that study is credible (Qi)	No. of procedures in each category* (Pi)	Probability of impact of mortality index based on maximum population base (Pi/4 385 597†)	Modified Qi based on impact probability (given by Qi × impact probability)
General/trauma 1	11·8	0·69	4 385 597	1·00	0·69
Trauma 1	13	0·76	129 514	0·03	0·02
General 1	13·4	0·79	4 256 083	0·97	0·76
Vascular 1	13·4	0·79	1 477 342	0·34	0·27
General/trauma 2	13·4	0·79	4 385 597	1·00	0·79
Trauma 2	13·4	0·79	129 514	0·03	0·02
Hip fracture	14	0·82	271 839	0·06	0·05
General 2	16	0·94	4 256 083	0·97	0·91
Orthopaedic	16	0·94	2 668 772	0·61	0·57
Vascular 2	16	0·94	1 477 342	0·34	0·32
General 3	16·4	0·96	4 256 083	0·97	0·94

Qi, quality score in the second column divided by 17. *Weighted national estimates from the Healthcare Cost and Utilization Project, Nationwide Inpatient Sample, 2008, Agency for Healthcare Research and Quality (AHRQ), based on data collected by individual states and provided to AHRQ by the states. †4 385 597 is the largest prevalence figure in the previous column.

Results

The search of MEDLINE, Embase, ERIC and the Cochrane Central Register of Controlled Trials identified 1169 unique abstracts (*Fig. 1*). The grey literature and BIOSIS Previews® search provided 94 abstracts, books and dissertations. All opinion papers, editorials and irrelevant abstracts (952 abstracts) were excluded. After applying the inclusion criteria, 100 relevant papers were examined, but eight papers from New York State and 29 from Europe were excluded; the remaining 63 papers examined the effect of the ACGME duty hours restriction. Based on the inclusion criteria, 53 of these were excluded as none looked at patient morbidity and mortality. Ten articles were finally included in this review[18–27] (*Table 1*).

Quantitative synthesis

The final pool consisted of ten studies with 19 data sets. *Tables S2* and *S3* (supporting information) show details of the studies, including criteria for mortality and morbidity.

Mortality data were available for eight studies[18–25] incorporating 11 data sets. There was a total of 730 648 patients: 385 639 before and 345 009 after introduction of the ACGME regulations. All were retrospective cohort studies. The mortality trials were deemed heterogeneous ($\tau^2 > 0$; $Q = 43{\cdot}77$; $P \leq 0{\cdot}001$), and not suitable for combination in one stratum using a fixed-effects model. The QE model was therefore used to pool the data[12]. The summary and individual ORs with 95 per cent confidence interval (c.i.) for each study are shown in *Fig. 2*. The summary OR for mortality pre- *versus* post-ACGME was 1·28 (95 per cent c.i. 0·94 to 1·73). This suggests that there was a 28 per cent greater odds of death before introduction of the regulations, although the confidence limits were wide and included 1. For comparison, a RE model was also examined (*Fig. S1*, supporting information). Interestingly, both the RE and QE models concurred in terms of direction of effect; only the RE model reached statistical significance, although there was a strong trend in the QE model. However, the RE model mortality results had a considerably smaller effect size estimate than the QE meta-analysis model (compare *Fig. 2* with *Fig. S1*, supporting information).

Morbidity data were available for seven studies[18,20,22,24–27] with eight data sets. There were 64 346 patients in total: 33 317 before and 31 029 after introduction of the ACGME working hours restriction. The morbidity trials were also deemed heterogeneous ($\tau^2 > 0$;

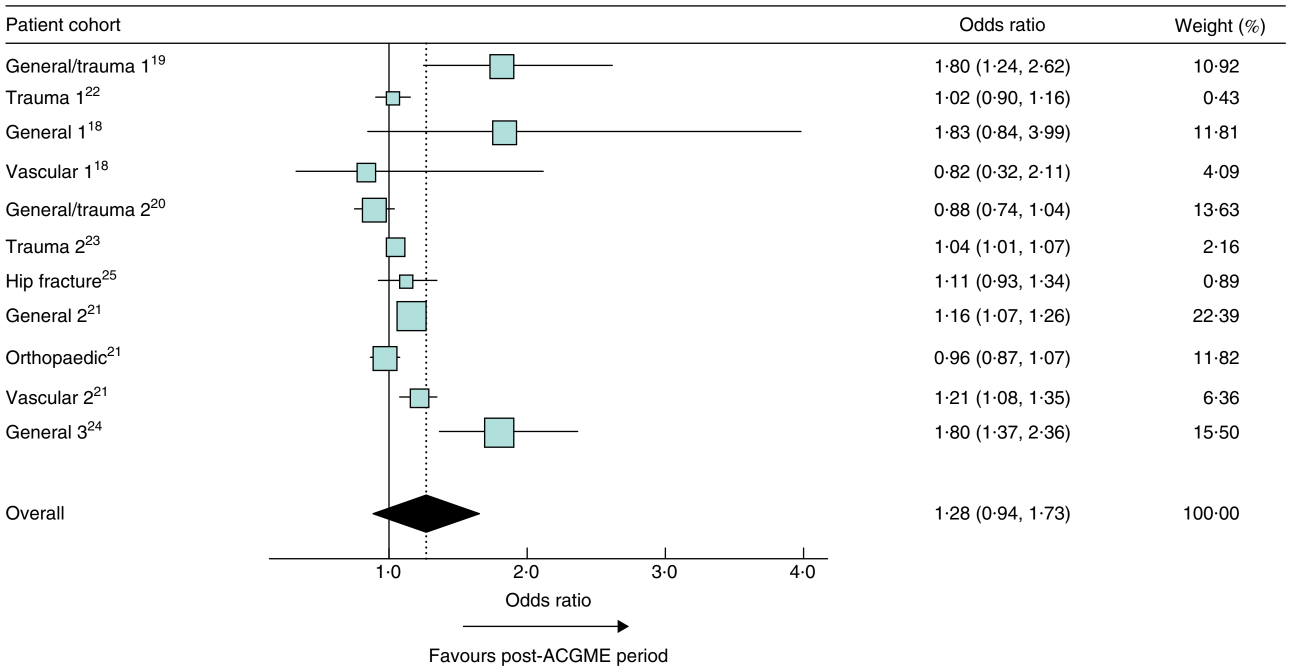

Fig. 2 Forest plot of the mortality odds ratio (OR) in individual studies and pooled mortality using the quality-effects (QE) model. ORs are shown with 95 per cent confidence intervals. The relative weights assigned by each model to each study are indicated by the size of the box. The QE weights take into account both the frequency of the surgical procedure and the quality of the study. ACGME, Accreditation Council of Graduate Medical Education

$Q = 237{\cdot}2$; $P < 0{\cdot}001$); the QE model was therefore used to pool the data, with a RE model being used for comparison. The summary and individual ORs and 95 per cent c.i. obtained using the QE model are shown in *Fig. 3*. The summary OR for morbidity pre- *versus* post-AGCME was 1·03 (95 per cent c.i. 0·67 to 1·57), suggesting no change. *Fig. S2* (supporting information) shows results for the RE model. Interestingly, both RE and QE models concurred in terms of magnitude of effect, but the direction of effect of the random model was opposite to that of the QE meta-analysis model (compare *Fig. 3* with *Fig. S2*, supporting information); however, the summary ORs were both close to 1.

Sensitivity analysis and publication bias for mortality studies

To determine how sensitive the findings were to changes in data included in the analysis, the effect of subgrouping of the various studies on the pooled results was analysed. The summary OR for the 11 data sets was determined after regrouping with respect to sample size (less than or equal to 17 000 patients, or more than 17 000 patients in the post-ACGME group), year of publication (2006 or earlier, 2007, and after 2007), type of surgery (general or other surgery) and study quality (quality score less than 8, or 8 or more). Associations were generally stronger for general surgery, more recent studies and higher-quality studies (*Table 2*). One-way sensitivity analyses with each trial individually removed suggested that no trial individually altered the pooled results significantly.

Analysis of the effect of potential unpublished studies or missed studies with negative result was undertaken to evaluate the robustness of the meta-analysis. The 95 per cent c.i. for the intercept (1·16) on Egger's regression for these studies included 0 (non-weighted) ($P = 0{\cdot}188$) suggesting no asymmetry, and thus no publication bias. Nevertheless, the funnel plot looked asymmetrical, and the QE pooled estimate deviated from the larger studies, implying that some degree of bias was likely (*Fig. 4*). There was thus evidence of asymmetry despite the non-significant Egger's regression results, suggesting that the larger studies represented either a less frequent procedure or were of poorer quality, or both.

Mortality before introduction of the working time regulations and implications of the quantitative synthesis

The mean mortality rate before introduction of the ACGME regulations was 4·2 (range 1·9–7·5) per cent. This wide variation in mortality rates was related to differences in the underlying surgical conditions and the variation in adherence to ACGME guidelines as well as

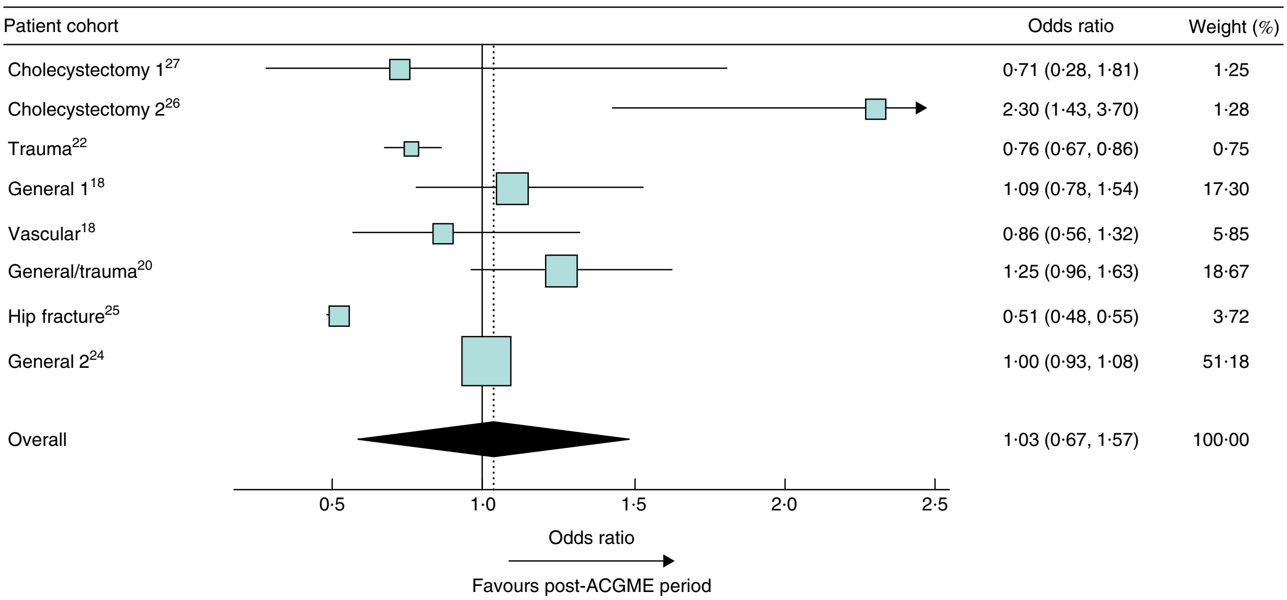

Fig. 3 Forest plot of the morbidity odds ratio (OR) in individual studies and pooled morbidity using the quality-effects (QE) model. ORs are shown with 95 per cent confidence intervals. The relative weights assigned by each model to each study are indicated by the size of the box. The QE weights take into account both the frequency of the surgical procedure and the quality of the study. ACGME, Accreditation Council of Graduate Medical Education

Table 2 Results of sensitivity analysis for mortality

Criteria	Model	Odds ratio
Type of surgery		
General	QE	1·18 (0·96, 1·48)
	RE	1·33 (1·02, 1·73)
Other	QE	1·03 (0·79, 1·36)
	RE	1·06 (0·99, 1·12)
Publication date		
2005–2006	QE	1·07 (0·77, 1·48)
	RE	1·24 (0·76, 2·04)
2007	QE	1·11 (0·97, 1·27)
	RE	1·09 (0·98, 1·20)
2009	QE	1·71 (0·96, 3·03)
	RE	1·24 (0·96, 1·59)
Size of post-ACGME group		
≤ 17 000	QE	1·32 (0·92, 1·90)
	RE	1·26 (0·96, 1·67)
> 17 000	QE	1·10 (0·97, 1·26)
	RE	1·09 (1·01, 1·17)
Quality score		
≥ 0·8	QE	1·19 (0·99, 1·44)
	RE	1·18 (1·03, 1·35)
< 0·8	QE	1·09 (0·89, 1·34)
	RE	1·06 (0·93, 1·21)

Values in parentheses are 95 per cent confidence intervals. QE, quality effects; RE, random effects; ACGME, Accreditation Council of Graduate Medical Education.

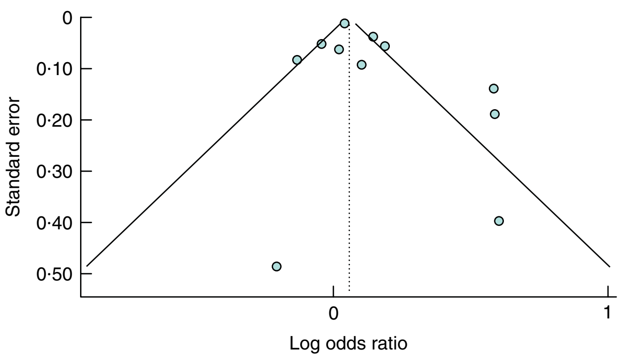

Fig. 4 Funnel plot demonstrating the possibility of asymmetry and thus publication (or other) bias. The central line represents the inverse variance weighted estimate used here for the sole purpose of judging symmetry

declining surgical mortality with time. In the quantitative synthesis the pooled pre- to post-ACGME OR was 1·28, representing a decrease in the odds of death of about 28 per cent after introduction of the ACGME guidelines. Part of this reduction in risk may be related to time period lags and simply reflect improving surgical mortality with time. However, there was no indication that moving to shorter working hours for surgeons would alter this trend unfavourably.

Discussion

The surgical community in North America disagrees with reducing duty hours[8] because of its anticipated effect on reducing continuity of care and minimizing educational opportunities for surgical trainees. However, despite these concerns, it seems that there will be further reductions as indicated in the recent ACGME task force recommendations[2]. The reason behind the further reduction in duty hours is the 2008 Institute of Medicine report, which raised concern that the 2003 ACGME duty hours restriction did not actually increase the amount of rest or sleep among residents[28]. This study has examined whether this translates, in addition, to mortality and morbidity concerns associated with long duty hours.

The total number of patients involved in the morbidity analysis was 64 346. Only one morbidity paper, which examined bile duct injuries and overall complication rates in patients undergoing laparoscopic cholecystectomy, found a statistically significant reduction in morbidity following the working hours reduction[26]. De Virgilio and colleagues[20] showed a decrease in morbidity in a surgery and trauma cohort but this was not statistically significant, and Kaafarani and co-workers[18] reported a non-significant reduction in morbidity among patients undergoing general surgery[13]. One data set showed no change in morbidity among patients having general surgery[24]. The remaining data sets showed worsening morbidity among vascular[18], trauma[22], orthopaedic[25] and general surgery[27] cohorts, but not all were statistically significant. Pooling of these studies using frequency of procedure weights did not suggest an adverse morbidity trend. This is an important finding as most of the literature linking fatigue to medical errors did not involve surgical patients treated by surgical residents. Indeed, one of the most cited studies on the subject of resident fatigue and medical errors is a randomized controlled trial comparing a traditional schedule with a schedule of reduced working hours, which found a reduction in non-preventable medical errors in the group of interns who worked fewer hours[9].

In contrast to the morbidity trend, several reports included in the present study found that the mortality rate was higher in surgical patients before the ACGME working hours restrictions[18,19,21–25]. The quantitative synthesis suggests that overall there was a 28 per cent decrease in the odds of death after the introduction of the ACGME regulations. Although this could be attributed to the decreasing mortality rate expected with time, it does reflect the fact that any decrease in continuity of care resulting from the working hours restriction did not reverse this trend. Eight papers containing 11 data sets with a total of 730 648 patients[13] examined mortality in surgical patients[18–25].

Only three data sets showed an increased mortality rate in the interval after introduction of the working hours restrictions, in orthopaedic[21], vascular[18] and a mixed general surgery/trauma cohort[20]. The remaining eight data sets showed a reduction in mortality, in particular all three general surgery cohorts[18,21,24].

The reduction in mortality demonstrated here could also be explained by advances in medicine. There have been improvements, for example, in intensive care that have been associated with mortality reductions[29]. Furthermore, the increasing preference for day surgery, with accompanying reductions in anaesthetic complications and nosocomial infections, and improved screening and early diagnosis (particularly for cardiovascular conditions and cancers), which result in elective rather than emergency surgical interventions, could have reduced adverse procedural outcomes in the period after the introduction of the ACGME regulations. These are just examples of the many factors that could similarly explain the reduction in mortality, which was not influenced solely by the reduction in working hours. What is clear, however, is that there was no demonstrable adverse trend in morbidity or mortality in the post-ACGME time period. It must nevertheless be kept in mind that all the included studies were observational, and confounded by the well recognized phenomenon that surgical outcomes improve over time. As working hours restrictions are 'late' developments in all the studies, an apparent 'benefit' is to be expected (even if they had no effect, or even an independent harmful effect that was smaller than the time trend effect). It is also possible that there was an independent harmful effect of working hours restrictions that was not detected owing to limitations in the design of the included studies. However, similar results were obtained in another study even when a multiple time series research design was used[30], lending credence to the lack-of-harm hypothesis.

The other concern surgeons have raised is whether the lack of an adverse trend is due to increased supervision, rather than reduced fatigue among residents; this remains a matter of debate. The Institute of Medicine report[28], as well as the recent recommendations of the ACGME task force, emphasized the importance of increasing supervision. Certainly the 2003 ACGME duty hours restriction aimed to increase physician attention to the importance of supervision in addition to reduced fatigue. In fact Privette and colleagues[24] found a reduction in mortality rate in a general surgery cohort after the introduction of reduced duty hours, but at the same time noted a doubling of attending surgeons' duty hours, suggesting that decreases in working hours for residents requires an increase in residents' supervision.

Every surgical procedure has a different volume, and so mortality or morbidity from a high-volume procedure makes up a greater bulk than that from a low-volume procedure. Even if the mortality rate after a low-volume procedure (such as hip replacement) is different from that after a high-volume one (for example general surgery), the impact on overall surgical mortality is closely linked to volume and needs to be factored into the meta-analysis weight[12]. Although the standard approach for handling heterogeneity between studies is to use the RE model, volume of the surgical procedure cannot be incorporated and so the RE model, in this respect, has no utility. Although this sort of weighting has been attempted in different ways by previous researchers[31,32], the present study used the QE model[12,33]. The methodology in the RE model is, in the authors' opinion, grossly flawed to the extent that, even in standard meta-analyses, the RE model estimate probably has no real interpretation[11,34]. The RE results are, however, noted for comparative purposes, especially as they reveal a significant effect in terms of the mortality endpoint, albeit of doubtful interpretation.

Finally, all papers were retrospective cohort studies looking simply at a category of patients and comparing their risk of death before and after the reduction in duty hours in teaching hospitals. There is a need for more rigorous research on medical errors in the surgical environment, which is a unique and different environment from that of other medical specialties. Even though the present data do not lend support to the contention that shorter working hours adversely affect mortality and morbidity, restricting duty hours may interrupt continuity of care. Lack of a direct effect of discontinuity of care on patient outcomes may be explained by the need for increased surgical supervision, which adds to the sense of dissatisfaction felt by staff surgeons; the opinion of surgical experts who are dissatisfied with patient care issues surrounding continuity of care[1] cannot simply be ignored.

Addendum

Two studies with mortality information were published after September 2009[35,36]. Addition of these two studies to the meta-analysis resulted in the same mortality odds ratio in the QE model (1·28) but a widening of the c.i. to 0·87 to 1·90.

Acknowledgements

M. H. Jamal and S. A. R. Doi contributed equally to this article. The authors thank Giovanna Badia from the Royal Victoria Hospital Library for help with the literature search.

Disclosure: The authors declare no conflict of interest.

References

1 Jamal MH, Rousseau MC, Hanna WC, Doi SA, Meterissian S, Snell L. Effect of the ACGME duty hours restrictions on surgical residents and faculty: a systematic review. *Acad Med* 2011; **86**: 34–42.

2 Nasca TJ, Day SH, Amis ES Jr; ACGME Duty Hour Task Force. The new recommendations on duty hours from the ACGME Task Force. *N Engl J Med* 2010; **363**: e3.

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

4 West D, Codispoti M, Graham T; Specialty Advisory Board in Cardiothoracic Surgery of The Royal College of Surgeons of Edinburgh. The European Working Time Directive and training in cardiothoracic surgery in the United Kingdom. *Surgeon* 2007; **5**: 81–85.

5 Lim E, Tsui S; Registrars and Consultant Cardiac Surgeons of Papworth Hospital 2003–2005. Impact of the European Working Time Directive on exposure to operative cardiac surgical training. *Eur J Cardiothorac Surg* 2006; **30**: 574–577.

6 Glomsaker TB, Søreide K. Surgical training and working time restriction. *Br J Surg* 2009; **96**: 329–330.

7 Maggiori L, Roupret M, Lefèvre JH. Workload and surgical training of residents in France: results of a national inquiry. *J Visc Surg* 2011; **148**: e141–147.

8 American College of Surgeons Task Force. Position of the American College of Surgeons on restrictions on resident work hours. *Bull Am Coll Surg* 2009; **94**: 11–18.

9 Landrigan CP, Rothschild JM, Cronin JW, Kaushal R, Burdick E, Katz JT *et al*. Effect of reducing interns' work hours on serious medical errors in intensive care units. *N Engl J Med* 2004; **351**: 1838–1848.

10 Harden RM, Grant J, Buckley G, Hart IR. Best evidence medical education. *Adv Health Sci Educ Theory Pract* 2000; **5**: 71–90.

11 Al Khalaf MM, Thalib L, Doi SA. Combining heterogenous studies using the random-effects model is a mistake and leads to inconclusive meta-analyses. *J Clin Epidemiol* 2011; **64**: 119–123.

12 Doi SA, Barendregt JJ, Mozurkewich EL. Meta-analysis of heterogenous clinical trials: an empirical example. *Contemp Clin Trials* 2011; **32**: 288–298.

13 Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F *et al*.; International Stroke Trial Collaborative Group; European Carotid Surgery Trial Collaborative Group. Evaluating non-randomised intervention studies. *Health Technol Assess* 2003; **7**: iii–x, 1–173.

14 Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998; **52**: 377–384.

15 Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological index for non-randomized studies (minors): development and validation of a new instrument. *ANZ J Surg* 2003; **73**: 712–716.

16 Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M *et al*. *The Newcastle–Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-analyses*. http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm [accessed 15 June 2007].

17 Concato J. Observational *versus* experimental studies: what's the evidence for a hierarchy? *NeuroRx* 2004; **1**: 341–347.

18 Kaafarani HM, Itani KM, Petersen LA, Thornby J, Berger DH. Does resident hours reduction have an impact on surgical outcomes? *J Surg Res* 2005; **126**: 167–171.

19 Malangoni MA, Como JJ, Mancuso C, Yowler CJ. Life after 80 hours: the impact of resident work hours mandates on trauma and emergency experience and work effort for senior residents and faculty. *J Trauma* 2005; **58**: 758–761.

20 de Virgilio C, Yaghoubian A, Lewis RJ, Stabile BE, Putnam BA. The 80-hour resident workweek does not adversely affect patient outcomes or resident education. *Curr Surg* 2006; **63**: 435–439.

21 Volpp KG, Rosen AK, Rosenbaum PR, Romano PS, Even-Shoshan O, Canamucio A *et al*. Mortality among patients in VA hospitals in the first 2 years following ACGME resident duty hour reform. *JAMA* 2007; **298**: 984–992.

22 Salim A, Teixeira PG, Chan L, Oncel D, Inaba K, Brown C *et al*. Impact of the 80-hour workweek on patient care at a level I trauma center. *Arch Surg* 2007; **142**: 708–712.

23 Morrison CA, Wyatt MM, Carrick MM. Impact of the 80-hour work week on mortality and morbidity in trauma patients: an analysis of the National Trauma Data Bank. *J Surg Res* 2009; **154**: 157–162.

24 Privette AR, Shackford SR, Osler T, Ratliff J, Sartorelli K, Hebert JC. Implementation of resident work hour restrictions is associated with a reduction in mortality and provider-related complications on the surgical service: a concurrent analysis of 14 610 patients. *Ann Surg* 2009; **250**: 316–321.

25 Browne JA, Cook C, Olson SA, Bolognesi MP. Resident duty-hour reform associated with increased morbidity following hip fracture. *J Bone Joint Surg Am* 2009; **91**: 2079–2085.

26 Yaghoubian A, Saltmarsh G, Rosing DK, Lewis RJ, Stabile BE, de Virgilio C. Decreased bile duct injury rate during laparoscopic cholecystectomy in the era of the 80-hour resident workweek. *Arch Surg* 2008; **143**: 847–851.

27 Naylor RA, Rege RV, Valentine RJ. Do resident duty hour restrictions reduce technical complications of emergency laparoscopic cholecystectomy? *J Am Coll Surg* 2005; **201**: 724–731.

28 Ulmer C, Miller-Wolman D, Johns MM (eds). *Resident Duty Hours: Enhancing Sleep, Supervision and Safety*. National Academies Press/Institute of Medicine: Washington, DC, 2008.

29 Rivers E, Nguyen B, Havstad S, Ressler J, Muzzin A, Knoblich B *et al*.; Early Goal-Directed Therapy

Collaborative Group. Early goal-directed therapy in the treatment of severe sepsis and septic shock. *N Engl J Med* 2001; **345**: 1368–1377.
30 Volpp KG, Rosen AK, Rosenbaum PR, Romano PS, Even-Shoshan O, Wang Y *et al*. Mortality among hospitalized Medicare beneficiaries in the first 2 years following ACGME resident duty hour reform. *JAMA* 2007; **298**: 975–983.
31 Saha S, Chant D, McGrath J. Meta-analyses of the incidence and prevalence of schizophrenia: conceptual and methodological issues. *Int J Methods Psychiatr Res* 2008; **17**: 55–61.
32 Nicholson WK, Robinson KA, Smallridge RC, Ladenson PW, Powe NR. Prevalence of postpartum thyroid dysfunction: a quantitative review. *Thyroid* 2006; **16**: 573–582.
33 Doi SA, Thalib L. A quality-effects model for meta-analysis. *Epidemiology* 2008; **19**: 94–100.
34 Doi SA. The case of the king's new clothes: there is no such thing as a random effects meta-analysis. *BMJ Rapid Response*. 29 April 2011; http://www.bmj.com/rapid-response/2011/11/03/case-Kings-new-clothes-there-no-such-thing-random-effects-meta-analysis [accessed 25 September 2011].
35 Helling TS, Kaswan S, Boccardo J, Bost JE. The effect of resident duty hour restriction on trauma center outcomes in teaching hospitals in the state of Pennsylvania. *J Trauma* 2010; **69**: 607–612.
36 Gopaldas RR, Chu D, Dao TK, Huh J, LeMaire SA, Coselli JS *et al*. Impact of ACGME work-hour restrictions on the outcomes of coronary artery bypass grafting in a cohort of 600 000 patients. *J Surg Res* 2010; **163**: 201–209.

Supporting information

Additional supporting information may be found in the online version of this article:

Appendix S1 Search strategy (Word document)

Appendix S2 Quality scoring (Word document)

Fig. S1 Mortality analysed in a random-effects model (Word document)

Fig. S2 Morbidity analysed in a random-effects model (Word document)

Table S1 Methodological quality scoring system (Word document)

Table S2 Mortality studies (Word document)

Table S3 Morbidity studies (Word document)

Please note: John Wiley & Sons Ltd is not responsible for the functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Commentary

Systematic review and meta-analysis of the effect of North American working hours restrictions on mortality and morbidity in surgical patients (*Br J Surg* 2012; 99: 336–344)

There has been much attention paid to restriction of surgeon trainee working hours, in both the USA and the countries of the European Union, in the hope that somehow patient safety would improve and trainees would be relieved of the presumed stress of long duty periods. It is important to realize that optimum patient care and safety should remain independent of the presence of surgical trainees. In other words, concern for, and initiatives to improve, patient outcomes and protect their welfare should apply across the continuum of healthcare delivery and not be confined to teaching institutions. Trainees should neither add to, nor detract from, this endeavour. Patient safeguards must interface with all healthcare providers. Reassuringly, in the present review, Jamal and colleagues, by culling the published literature, have demonstrated that patient outcomes (mortality and morbidity) – perhaps as surrogates for safety – have remained unaffected, or even improved, very likely irrespective of adjustments to trainees' working hours.

What may be more relevant, as the American College of Surgeons has pointed out in its response to the Institute of Medicine's report on working hours restriction[1], is whether working hours reductions have affected the quality of trainee education and the ability of these future surgeons to function as practitioners in a stressful and demanding profession. The present study did not address this, but the authors acknowledged the paucity of valid scientific evidence concerning working hours issues, a finding similar to that reported earlier this year by Moonesinghe and co-authors[2]. Now that we can be somewhat assured that strategies to optimize patient outcomes and safety have superseded training paradigms, let us focus on the essence of the duty hour restriction controversy: can we adequately train our future surgeons in the amount of time allocated by regulatory agencies? The need for more and better designed studies to resolve this question is clear. However, whether solid scientific evidence will be enough is a matter of conjecture.

T. S. Helling
Department of Surgery, University of Mississippi Medical Center, 2500 North State Street, Jackson, Mississippi 39216, USA
(e-mail: thelling@umc.edu)
DOI: 10.1002/bjs.8655

Disclosure

The author declares no conflict of interest.

References

1 Britt LD, Sachdeva AK, Healy GB, Whalen TV, Blair PG; Members of ACS Task Force on Resident Duty Hours. Resident duty hours in surgery for ensuring patient safety, providing optimum resident education and training, and promoting resident well-being: a response from the American College of Surgeons to the report of the Institute of Medicine, 'Resident Duty Hours: Enhancing Sleep, Supervision, and Safety'. *Surgery* 2009; **146**: 398–409.
2 Moonesinghe SR, Lowery J, Shahi N, Millen A, Beard JD. Impact of reduction in working hours for doctors in training on postgraduate medical education and patients' outcomes: systematic review. *BMJ* 2011; **342**: d1580.