ISSN 2379-4046

EMERGENCY MEDICINE
Open Journal

Openventio PUBLISHERS

Systematic Review

Long Working Hour Related Medical Errors and Patient Outcomes among Physicians: A Systematic Review

Wei-Yo Chao, MD[1*]; Shyr-Chyr Chen, PhD[2]; Emma Geerdes, MD[3]; Jo-Ching Hsiung, MD[1]; Cheng-Han Li, MD[1]; Chien-Chang Lee, PhD[2]; Haruna Tanoue, MD[4]; Mai Yamazaki, MD[4]

[1]*School of Medicine, College of Medicine, National Taiwan University, Taiwan (R.O.C.)*
[2]*Department of Emergency Medicine, National Taiwan University Hospital, and National Taiwan University, College of Medicine, Taipei, Taiwan (R.O.C.)*
[3]*School of Medicine, College of Medicine, Erasmus University Rotterdam, The Netherlands*
[4]*School of Medicine, College of Medicine, Osaka University, Japan*

[*]***Corresponding author***
Wei-Yo Chao, MD
School of Medicine, College of Medicine, National Taiwan University, Taiwan (R.O.C.); Tel. +886978883791; E-mail: b04401109@ntu.edu.tw

Article information
Received: October 4th, 2018; **Revised:** December 12th, 2018; **Accepted:** December 12th, 2018; **Published:** January 8th, 2019

Cite this article
Chao W-Y, Chen S-C, Geerdes E, et al. Long working hour related medical errors and patient outcomes among physicians: A systematic review. *Emerg Med Open J.* 2019; 4(1): 25-35. doi: 10.17140/EMOJ-4-151

ABSTRACT

Aim

Inspired by the famous Libby Zion case in 1984, which revealed the underlying flaw in the medical system at the time——the exploitation of junior medical staff and the inadequate surveillanceby attending physicians, authors of this systemic review aim to investigate the potential consequence of overworking in medical scenarios.

Methods

Excerpta medica database (EMBASE) and PubMed had been systematically searched, bibliographies of relevant studies additionally reviewed. Four cohorts and three randomized controlled trials were selected and quality-assessed with Newcastle-Ottawa Scale and the Cochrane risk of bias tool, respectively. Literature were extracted and discussed.

Results

Two randomized controlled trials have concluded that residents or interns with longer consecutive working hours per shift in the division of internal medicineare prone to more medical errors. One cohort study has shown a significant association between longer weekly working hours and worsened quality of patient care (intensive care unit (ICU) transfer rates and in-hospital mortality) in the setting of internal medicine. However, none of the 3 studies which were conducted in surgical departments suggests a further restriction on weekly working hours or shift length.

Conclusion

The specialty-specific policy is recommended based on our between-study comparison. Specifically, with respect to the incidence of medical errors or adverse events, weekly hours or shift length are weakly recommended to be regulated in the department of internal medicine, but there is no recommendation to surgical departments. However, one must consider all respects of the impact brought by any alternation of working policy before the actual implementation.

Keywords

Physicians; Residents; Interns; Duty hours; Working hours; Accreditation council for graduate medical education (ACGME); Complication; Medical error; Adverse events; Patient Safety.

INTRODUCTION

Research by Grober and Bohnen[1] defined the medical error as "an act of omission or commission in planning or execution that contributes or could contribute to an unintended result." Similarly, the definition of an adverse event is an "unintended injury to patients caused by medical management (rather than the underlying condition of the patient) that results in measureable disability, prolonged hospitalization, or both." Not all medical errors necessarily result in adverse events and vice versa. Any occurrence of a medical error or a preventable adverse event may reveal underlying problems in any section before the final execution of a medical order. Overworking, for instance, is regarded as the culprit for the incidence of medical errors or adverse events. One prospective 2-session within-subject study shows that residents with heavy night-call duty and placebo drug ingestion have cognitive performance comparable to residents following light-call duty with alcohol intake, blood-alcohol concentration falling in the range of

0.04~0.05 g%.[2] An other comparative researchindicated interns with significantly more weekly working-hours and significantly less sleeping-hours had more attentional failure.[3] Furthermore, having discovered more recent articles, we aim to make an update of the concerned topic by conducting a literature systematic review, investigating whether long working hours and related factors will lead to higher medical-error rates or worsened patient outcomes.

METHODOLOGY

Data Sources and Searches

This systematic review is in accordance with the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines.[4] Ethical Committee approval is not required. The primary author (C W-Y) searched excerpta medica database (EMBASE) and PubMed in January 2018. The searching strategies of both EMBASE and PubMed are attached in the Appendix. The author also manually reviewed bibliography of relevant studies and other systematic reviews to ensure the completeness the literature search.

Study Selection

The primary author (C W-Y) firstly screened titles and abstracts from primary records. Duplicates, unrelated studies and studies without full-acccess were excluded. The inclusion criteria are as follows: the population should be physicians, residents, house-staff, or interns, and the comparison of working schedules between the intervention and the control groups should be demonstrated in the study. Outcomes encompass the incidence of medical errors, adverse events, quality indicators of healthcare, or post-operation complications. Outcomes such as the records of errors in surgical simulations or training, cognitive performance, and records of attentional failure are excluded from consideration.

Within articles extracted from the database, many articles investigating the effectiveness of either the Code 405 in New York State, or 2003, 2011 accreditation council for graduate medical education (ACGME) act of regulation on the duty hour in the U.S. were retrieved. Among the primarily retrieved record from the database, there are many studies investigating the effectiveness of either the Code 405 in New York State, or 2003, 2011 accreditation council for graduate medical education (ACGME) act of regulation on the duty hour in the US., demonstrating crude data of mortality or complication rates. Even so, these studies without presentation the difference of working-hours or the compliance of each resident training program to their investigated policies is not accepted. Systematic reviews, letters, editorials, and surveys are excluded. Only cohort studies and randomized controlled trials are eventually included. Studies sourced through the bibliography-review were documented in the *Additional records* of our *PRISMA* 2009 Flow diagram.

Quality Assessment

The primary author (C W-Y) adopted the Newcastle-Ottawa quality assessment and The Cochrane risk of bias tool (modified by Dr. Lee C-C) to conduct the quality assessment for cohort studies and randomized controlled trials, respectively.

Information obtained from all studies includes (1.) Location where studies were conducted (2.) Investigated specialties (3.) Patient numbers (4.) Schedules of the intervention groups and the control groups (5.) Study types and designs (6.) Periods between the beginning and the end of each research (7.)Characteristics of patients and participants at baseline (9.) p value presentation of schedule difference (10.) Postgraduate trainees' involvement in patient contact (11.) Supervision involvement (12.) Kappa statistics to assess discrepancy of review's opinions on incident reports (13.) Methods used to document data of working hour, medical errors, adverse events, and patient outcomes to determine the validity of both the exposure and the outcome.

Terms to assess the risk of bias for cohort studies include (1.) Representatives to the exposed cohort (2.) Selection of the non-exposed cohort (3.) Ascertainment of exposure (4.) Demonstration that outcome of interest was not present at the start of the study (5.) Comparability (6.) Assessment of outcome (7.) Was follow-up long enough for outcomes to occur, and (8.) Adequacy of follow up of cohorts. Answers to each question are ranked by 0, 1, or 2 stars and are categorized into "Selection, Comparability, Exposure, and Outcome."

Terms to assess the risk of bias for randomized controlled trials include (1.) Random sequence generation (2.) Allocation concealment (3.) Blinding of a participants and personnel (4.) Blinding of outcome assessment (5.) Incomplete outcome and (6.) Selective reporting. Each question is assessed as "High Risk", "Low Risk", or "Unclear" based on the methodology of collected studies.

Discrepancies were resolved through discussions between the author and the third investigator, Dr. Lee C-C. The recommendation of this article was based on the evaluation of the strength and the weakness of each included studies with grading of recommendations assessment, development and evaluation (GRADE) guideline. Additionally, the critical appraisal skills programme (CASP) checklist for systematic reviews[5] are referenced to comment on one another systematic review, presented in the discussion. Literature extraction is presented in Tables 1-7.

RESULT

Study Selection and Characteristics

Numbers of excluded articles are reported with reasons in the PRISMA flow diagram (Figure 1). Primary records include 502 articles and 25 from bibliography review. Following the deletion of duplicates and the abstract screening, 122 eligible articles undergo the process of full-text review. With the inclusion criteria matched, 4 retrospective cohort studies and 3 randomized controlled trials were selected. Among the 7 studies, 3 studies were conducted in surgery departments with the rest 4 belonging to the category of internal medicine. Six studies were conducted in the U.S. and one in the U.K., and there are 6 single-center studies and 1 multi-center studies (Tables 1 and 2).

Figure 1. *Flowchart of the Literature Search*

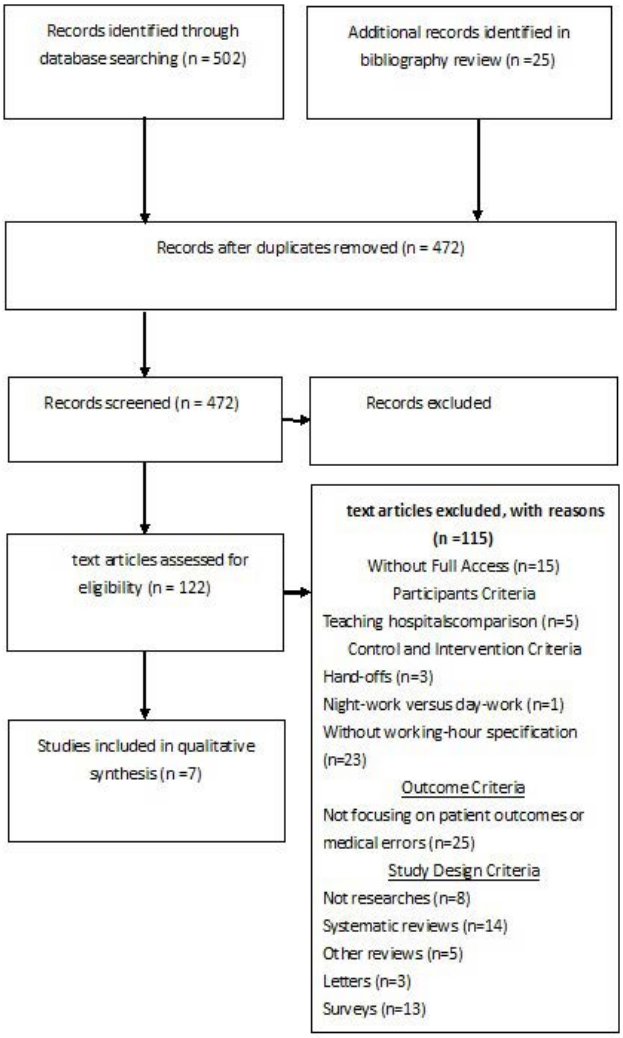

Table 1. *Background Settings (Locations and the Specialty); Schedule or Policy Settings; and Numbers of Patients*

			Patient Number		Intervention Schedule		Control Schedule	
Sources	**Location**	**Specialty**	**Intervention**	**Control**	**Weekly working hours**	**Shift length (Hours)**	**Weekly working hours**	**Shift length (Hours)**
Anderson et al[10]	The U.S.	Surgery	N=198	N=185	<80	On call: 26-28 Not on call: 10-14	<80	Interns:16 Residents:24
Bhavsar et al[13]	The U.S.	Internal Medicine	N=431	N=572	<80	Interns&:35 Residents:35	<80	Interns:30 Residents:14
Ouyang et al[12]	The U.S.	Internal Medicine	N=1,411	N=2,039	≥80	Not specified	<80	Not specified
Yahoubian et al[15]	The U.S.	Surgery	N=878 (-A) N=2,522 (-C)	N=708 (-A) N=386 (-C)	Not specified	<16	Not specified	>16
Cappuccio et al[9]	The U.K.	Internal Medicine	N=238	N=244	43.2	9-11	52.4	9.9-12.5
Bilimoria et al[11]	The U.S.	Surgery	N=65,849	N=72,842	<80	No restriction	<80	Interns≤16 Residents≤24
Landrigan et al[14]	The U.S.	Internal Medicine	N=227	N=354	60-63	7-16	77-81	7-29

Table 2. *Study Types; Designs; Periods of the Research; p value of Schedule Difference; Specific Patient Groups; Characteristics Demonstration of the Participants and the Patients at Baseline; PGs'(Postgraduate Trainees') Involvement in Patient Contact; whether Medical Action was Under team-Supervision; Kappa Statistic for Assessment of Discrepancy among Incident Report Reviewers' Opinions; Data Collection of Working Hours, Medical Errors and Patient Outcome*

Sources	Study type	Design	Periods of the research	*p* value of schedule difference	Specific patient groups	Characteristics demonstration of patients/participants at baseline
Anderson et al[10]	Retrospective cohort study	Crossover	2 years	No	No	Not specified / Not specified
Bhavsar et al[13]	Retrospective cohort study	Crossover	2 years	No	Yes	Not specified / Yes
Ouyang et al[12]	Retrospective cohort study	Parallel	1 year	No	No	Not specified / Yes
Yahoubian et al[15]	Retrospective cohort study	Parallel	7 years	No	Yes	Yes / Yes
Cappuccio et al[9]	Retrospective cohort study	Parallel	12 weeks	Yes	Yes	Yes / Yes
Bilimoria et al[11]	Randomized controlled trial	Parallel	1 year	No	Yes	Yes / Yes
Landrigan et al[14]	Randomized controlled trial	Parallel	1 year	No	No	Yes / Yes
Sources	**PGs' involvement**	**Under supervision**	**Kappa statistics**	**Data collection of working hours**	**Data collection of medical errors**	**Data collection of AEs or patient outcomes**
Anderson et al[10]	81.4%	Not specified	Not specified	Not specified	Self-report	Self-report
Bhavsar et al[13]	100%	Yes	Not specified	Not specified	Not applicable	Electronic records
Ouyang et al[12]	100%	Yes	Not specified	Electronic records	Not applicable	Electronic records
Yahoubian et al[15]	100%	Yes	Not specified	Not specified	Not applicable	Electronic records
Cappuccio et al[9]	100%	Not specified	0.8	Self-report	Electronic records	Electronic records
Bilimoria et al[11]	Not specified	Not specified	Not specified	Not specified	Not applicable	Electronic records
Landrigan et al[14]	100%	Yes	0.8-0.9	Self-report	Electronic records	Electronic records

Risk of Bias Assessment of Cohort Studies

All patients from the exposed group in 4 cohort studies truly represent the community of interests. Non-exposed participants were drawn from the identical community as the exposed. Exposure ascertainment is evaluated by the method to obtain figures of the working-hour. Only in the Ouyang et al[6] the electronic record was used to document the time during work: the time difference between the first and the last medical action made. By contrast, the compliance rate was unclear in other cohorts. Violation of the duty hour policy results in exposure validity unless the record was secured by electronic data. Similarly, how the data of the outcome was retrieved in a study affect the outcome validity. Participant residents were unable to be independently blinded from the knowledge to the exposure state. Besides, Anderson et al[7] collected data *via* self-reporting when there is record linage in Ouyang et al,[6] Yahoubian et al[8] and Bhavsar et al[9] The follow-up was long enough for the outcome to be visible in the 4 cohorts. On the other hand, no statement of the adequacy of the follow-up was noted. Patient outcomes may be associated with the pre-admission condition such as inherent systemic or chronic diseases. However, the outcome of interests at the entry endpoint was not discussed in these studies. Primary exposures are either shift-length or weekly working hours. Other variables in the intervention design may also include the length for clinical education, hand-over between shifts, sleeping hours as expected to be associated with a change of schedule (Tables 3.1 and 3.2).

Table 3.1. *Quality Assessment for Cohort Based on the Newcastle-Ottawa Scale (NOS)*

Sources	Selection	Comparability	Exposure	*Outcome*	*Total NOS Stars*
Anderson et al[10]	***	**		*	6; High quality
Bhavsar et al[13]	**	**		***	7; High quality
Ouyang et al[12]	***	*	*	**	7; High quality
Yahoubian et al[15]	**	*		**	5; Medium quality

Table 3.2. *Quality Assessment for Cohort Based on the Newcastle-Ottawa Scale*

Sources	Representativeness of the Exposed cohort	Selection of the non-exposed cohort	Ascertainment of exposure	*Demonstration that outcome of interest was not presented at start of study (medical errors/patient outcomes)*
Anderson et al[10]	True representative of the average resident(including interns) in the community	Drawn from the same community as the exposed cohort	Unclear	Not applicable / Unclear
Bhavsar et al[13]	True representative of the average resident(including interns) in the community	Drawn from the same community as the exposed cohort	Unclear	Not applicable / Unclear
Ouyang et al[12]	True representative of the average resident(including interns) in the community	Drawn from the same community as the exposed cohort	Secure records	Not applicable / Unclear
Yahoubian et al[15]	True representative of the average resident(including interns) in the community	Drawn from the same community as the exposed cohort	Unclear	Not applicable / Unclear
Sources	**Comparability**	**Assessment of outcome**	**Was follow-up long enough for outcomes to occur?**	**Adequacy of follow up of cohorts**
Anderson et al[10]	Study controls for call shift length, additional controlled factor is time off between shifts or after 24-hour shifts	Self-report Blinding: Not applicable	Yes	No statement
Bhavsar et al[13]	Study controls for call shift length, additional controlled factor is time off between shifts or after 24-hour shifts	Record linage Blinding :Not applicable	Yes	No statement
Ouyang et al[12]	Study controls for call shift length	Record linage Blinding: Not applicable	Yes	No statement
Yahoubian et al[15]	Study controls for call shift length	Record linage Blinding: Not applicable	Yes	No statement

Risk of Bias Assessment of Randomized Controlled Trials

With respect to random sequence generation, high risk of bias in Cappuccio et al[9] is attributed to the lack of interest in participation of interns in the intervention arm due to the physical and mental exhaustion induced by the rapid change of circadian rhythm. As a result, the selection bias existed during the research. Secondly, the allocation concealment was attained in all studies. Blinding of the participants, however, was not feasible to be attained. On the other hand, following the participant interns' schedules, the observers (the person who observes the participant interns and records down each occurrence of medical errors or adverse events) in Landrigan et al[11] were not blinded. The medical record was further reviewed by two blinded independent reviewers. Results of most studies were assessed manually by reviewers. However, only Cappuccio et al[9] andLandrigan et al[11] provided kappa statistics to evaluate the discrepancy in the preliminary reviewing process, so-called inter-rater reliability. Disagreements were resolved by either re-discussion or seeking the third opinion in studies selected within the present review.

Both Cappuccio et al[10] and Bilimoria et al[12] used intention-to-treat analysis to address missing data, while Landrigan et al[11] didn't provide any information on how missing data were addressed; however, Landrigan et al[11] did note that only one intern was observed at a time because there were no enough observers, missing some dataevenly in both the intervention and the control arms. (Table 4).

Table 4. *Quality Assessment for Randomized Controlled Trial*

Sources	Random Sequence Generation	Allocation Concealment	Blinding of Participants and Personnel
Cappuccio et al[9]	High (due to working schedule design, some interns were reluctant to join.)	Low	High
Bilimoria et al[11]	Low	Low	High
Landrigan et al[14]	Unclear	Low	High
Sources	**Blinding of Outcome Assessment**	**Incomplete Outcome**	**Selective Reporting**
Cappuccio et al[9]	Low	Low (intension-to-treat analysis)	Unclear
Bilimoria et al[11]	Low	Low (intension-to-treat analysis)	Unclear
Landrigan et al[14]	Low (Un-blinded observers; blinded medical record reviewers)	Unclear	Unclear

Emerg Med Open J. 2019; 4(1): 25-35. doi: 10.17140/EMOJ-4-151

Results of Individual Studies

Bilimoria et al[12] and Anderson et al[7] adopted the identical schedule policy in accordance with 2011 ACGME reform and the flexibility in duty hour requirements for surgical trainees (FIRST) trial, which waived maximum shift-length and transitional periods between shifts——the reform limited weekly working hours to less than 80, shift-length to less than 24 to 28 hours for residents (4 hours for transition) and to less than 16 hours for interns, the transitional periods between shifts exceeding at least 8 to 10 hours and at least 14 hours after 24-hour on-call shifts. A randomized controlled trial owned profound evidence: Bilimoria et al[12] affirms that there was no significant difference in the in-hospital complication rate, or the rate of the 30-day post-operative mortality or serious complication.

Anderson et al[7] shows the insignificant difference of the incidence of medical errors in surgical departments of the university of California Davis Medical Center between the group following the FIRST policy and the other of the standard policy, thus no association is indicated. By contrast, significantly more "total adverse events," including unpreventable adverse events (the unavoidable injury resulting from appropriate medical care), were observed in the FIRST trial than in the group of the standard policy.

In Yahoubian et al,[8] surgeries were stratified according to the starting time into day-time surgery (from 6 A.M. to 10 P.M.) and night-time surgery (10 P.M. to 6 A.M.). The night-time surgeries were conducted by residents working beyond 16 hours. postgraduate year (PGY) levels were fairly distributed between groups and the operation time were similar as well. A policy was also implemented: Residents who were on the call was not allowed to operate after 6 A.M. in the next morning, thus distinguishing the night-time group from the day-time group. Between-group differences of complication rates were both insignificant in Yahoubian et al.[8]

In Cappuccio et al[10] the intervention group followed the European Working Time Directive (the EWTD) schedule, working 43.2 hours on average, while the control followed the traditional working hour in the U.K., working 52.4 hours in a week on average. Results demonstrated the significant between-group difference of weekly hours of work and a higher rate of intercepted adverse events (p=0.002) in the traditional working group compared to the group of EWTD schedule (Tables 5, 6.1 and 6.2).

Table 5. *Results of Medical Error Rates Across Studies*

	Medical Errors	
Sources	**Selection**	**Comparability**
Anderson et al[10]	Total errors (p=0.596)	
	n=488	n=472
	Preventable adverse events (p=0.618)	
	3.9 % of all operations during the period	3.7% of all operations during the period
	Cognitive errors (p=0.360)	
	n=114	n=115
	Judgement errors (p=0.196)	
	n=95	n=101
	Technical errors (p=0.196)	
	n=95	n=101
	Team failure errors (p=0.719)	
	n=64	n=63
Cappuccio et al[9]	Overall Medical errors (p=0.006)	
	27.6 medical errors happened in 1,000 patient-days	41 medical errors happened in 1,000 patient-days
Landrigan et al[14]	Serious medical errors (p< 0.001)	
	100.1 medical errors per 1,000 patient-days	132 medical errors per 1,000 patient-days
	Preventable adverse events (p=0.21)	
	16.5 medical errors per patient-days	20.9 medical errors per 1,000 patient-days
	Intercepted serious errors (p=0.02)	
	55 medical errors per 1,000 patient-days	70 medical errors per 1,000 patient-days
	Nonintercepted serious errors (p<0.001)	
	28.6 medical errors per 1,000 patient-days	44.8 medical errors per 1,000 patient-days
	Serious medication errors (p=0.03)	
	82.5 medical errors per 1,000 patient-days	99.7 medical errors per 1,000 patient-days
	Serious procedural errors (p=0.34)	
	6.6 medical errors per 1,000 patient-days	8.5 medical errors per 1,000 patient-days
	Serious diagnostic errors (p<0.001)	
	3.3 medical errors per 1,000 patient-days	18.6 medical errors per 1,000 patient-days

Table 6.1. *Results of Patient Outcomes (Adverse Events and Complication Rates) Across Studies*

	Adverse Events or Complication Rates	
Sources	**Intervention**	**Control**
Anderson et al[10]	Total adverse events (p=0.026)	
	6.6% of all operations during the period	7.8% of all operations during the period
	Preventable adverse events (p=0.618)	
	3.9 % of all operations during the period	3.7% of all operations during the period
Bhavsar et al[13]	In-hospital adverse events (unadjusted)	
	Congestive heart failure/ Pulmonary edema	OR=1.14, 95% CI: [0.59 to 2.21], p=0.694
	Cardiogenic shock	OR=1.06, 95% CI: [0.57 to 1.96], p=0.85
	Cardiac arrest	OR=0.77, 95% CI: [0.42 to 1.43], p=0.409
	In-hospital mortality (unadjusted)	OR=0.65, 95% CI: [0.32 to 1.32], p=0.237
	MACE at 6 months following discharge	
	MACE at 6 months following discharge (unadjusted)	OR=0.76, 95% CI: [0.50 to 1.18], p=0.226
	MACE at 6 months following discharge (adjusted)	OR=1.29, 95% CI: [0.76 to 2.20], p=0.35
Ouyang et al[12]	Composite endpoint	OR=1.21, 95% CI: [1.02 to 1.43], p=0.027
	Transfer to ICU	OR=1.62, 95% CI: [1.13 to 2.32], p=0.008
	30-Day readmission	OR=1.07, 95% CI: [0.88 to 1.29], p=0.513
	In-hospital mortality	OR=1.44, 95% CI: [0.99 to 2.10], p=0.057
Yahoubian et al[15]	Overall complication rate (Appendectomy) (p=0.1)	
	1%	2%
	Overall complication rate (Cholecystectomy) (p=0.9)	
	2.3%	2.2%
Cappuccio et al[9]	Preventable adverse events (p=0.68)	
	1.6 medical errors happened in 1,000 patient-days	2.2 medical errors happened in 1,000 patient-days
	Intercepted potential adverse events (p=0.002)	
	3 medical errors happened in 2467 patient-days	16 medical errors happened in 2315 patient-days
	Non-intercepted potential adverse events (p=0.067)	
	41 medical errors happened in 2467 patient-days	56 medical errors happened in 2315 patient-days
Bilimoria et al[11]	Any complications	OR=0.94 (unadjusted), 95% CI: [0.84 to 1.06]
		OR=0.96 (adjusted), 95% CI: [0.89 to 1.04]
	Serious complications	OR=0.96 (unadjusted), 95% CI: [0.86 to 1.06]
		OR=0.96 (adjusted), 95% CI: [0.90 to 1.04]
	30-day rate of post-operative death or serious complications	OR=0.96 (unadjusted), 95% CI: [0.87 to 1.06]
		OR=0.96 (adjusted), 95% CI: [0.95 to 1.04]
Landrigan et al[14]	Preventable adverse events (p= 0.21)	
	On average 16.5 medical errors per 1,000 patient-day	On average 20.9 medical errors per 1,000 patient-day

Table 6.2. *Results of Patient Outcomes (Mortality) Across Studies*

	Mortality	
Sources	**Intervention**	**Control**
Anderson et al[10]	In hospital death (p=0.479)	
	12.6% in-hospital death rate	15.7% in hospital death rate
Bhavsar et al[13]	In-hospital mortality(unadjusted)	OR=0.65, 95% CI: [0.32 to 1.32], p=0.237
	In-hospital mortality(adjusted)	OR=0.47, 95% CI: [0.18 to 1.20], p=0.11
	Mortality at 6 months following discharge (unadjusted)	OR=0.40, 95% CI: [0.22 to 0.71], p=0.002
	Mortality at 6 months following discharge (adjusted)	OR=0.53, 95% CI: [0.28 to 0.99], p=0.05
Ouyang et al[12]	In-hospital mortality	OR=1.44, 95% CI: [0.99 to 2.10], p=0.057
Yahoubian et al[15]	Not eligible (casenumber=1, out of 1,587 patients in appendectomy)	
	Not eligible (case number=1, out of 1,666 patients in cholecystectomy)	
Cappuccio et al[9]	In-hospital mortality	OR=1.13 (age-adjusted), 95% CI : [0.88 to 1.44] , p=0.340
Bilimoria et al[11]	2 Death rate within 30 days after operations	OR=0.95 (Adjusted), 95% CI : [0.83 to 1.10]
Landrigan et al[14]	Not applicable	

In Landrigan et al[11] a trial of working schedule was performed: shift-length ranges from 7 to 16 hours, weekly working hours ranging from 60 to 63 in the intervention group. On the other hand, with the restriction lifted, the shift-length of the control group exceeded 29 hours and weekly working hours reached up to 77 to 81 hours during the trail. Results show insignificance between groups regarding the medication-error, procedural-error, and intercepted medical errors. However, serious medical error, the diagnostic medical error, and non-intercepted medical error in the intervention was significantly less than those of the control group. The incidence of preventable adverse events was not significantly different in Anderson et al,[7] Cappuccio et al,[10] and Landrigan et al.[11]

In Ouyang et al[6] significantly higher intensive care unit (ICU) transfer rate, the incidence of 30-day readmission, and more in-hospital death was observed in 40% of the patients taken care of by medical teams with at least one resident working over 80 hours per week, compared to the rest 60% who were taken care of by residents working less than 80 hours per week.

Bhavsar et al[9] a retrospective cohort study, investigates if the implementation of the 2003 ACGME policy changed the patient outcomes in the inpatient cardiology service. In the era prior to the policy implementation, all residents and interns had worked 35 hours overnight per shift. Within the 2-years, after the policy was implemented, on the other hand, interns were on-call from 7 a.m., however, to no more than 1 p.m. the next day while residents abided by the day-float system (14 hours at work in the daytime). Results of the study show in-hospital mortality rates were insignificantly associated with working schedule difference while the mortality rate at the 6th month of the follow-up is significantly higher among patients taken care of by clinicians working longer hours per week. Nevertheless, the mortality at the 6th month seems unlikely to be highly related to the quality of inpatient care because of causes of the death were unclear from the contexts.

An unexpected result is indicated: the weekly time on the arranged working schedule in the Cappuccio et al[9] a U.K. based study, was far too short compared to other studies conducted in the United States, though a significant between-group difference of outcomes of medical error rates was revealed in Cappuccio et al[9] while most studies in the U.S. was not. Nevertheless, since the shift length is similar between researches in the U.K. and the U.S., it is reasonable to expect that other factors such as the weekly hours at work or sleeping hours per week might play a role in the introduction of such a diverse result.

DISCUSSION

According to the present review, of no significance is one generalized suggestion for the working schedule to all specialties due to their different needs in the arrangement of staff-on-board. Therefore, to conclude (Table 7), weekly working hours and shift-length less than 80 and 16 hours, respectively, are likely to the reduce medical error rate and optimize the patient outcomes in the division of internal medicine. On the other hand, researches concerning the surgical division reveal duty hour regulation on residents might not be associated with the incidence of medical errors, or affect patient outcomes, a similar result in Amiran et al with the Karolinska scale and other indicators concerning laparoscopic skills utilized to observe interns, residents, and attending surgeons before and after a 17-hour night shift.

In Fletcher et al[13] a literature systematic review, all studies were searched systematically, appraised critically with searching the strategy provided in the context, strength and weakness of each study discussed in detail. Similarly, no definite conclusion was attained due to the diverse settings across studies.

There are multiple explanations of the phenomena. Gonzalo et al[18] points out that medical error rates or physician-patient satisfaction survey concerning the life quality after the change of working schedule might be more sensitive to the working hour schedule change instead of in-hospital mortality. Additionally, the indirectness, unclear ratios of compliance to the designated schedule for interns[21] or the unknown proportion of residents' involvement in each surgery or patient care,[10] may render the result invalid. Other similar concerns include the lack of significant difference between schedules of the intervention and that of the control group and the uncontrolled confounders such as resident-to-bed ratio, the length of surgery time, the sleeping hours, or the discontinued patient care.

Table 7. *Summary with the Quality of Evidence Abbreviation.:AEs, adverse events*

Sources	Specialty	Variables applied in the study		*Conclusion -Medical errors*	Conclusion -AEs & Patient outcomes	Quality of evidence
		Shift length	Weekly working hours			
Anderson et al[10]	Surgery	√		Insignificant association	Insignificant association	Medium
Bhavsar et al[13]	Internal Medicine	√		Not applicable	Insignificant association	Medium
Ouyang et al[12]	Internal Medicine	Not specified	√	Not applicable	pro low weekly working hours	Medium
Yahoubian et al[15]	Surgery	√	Not specified	Not applicable	Insignificant	Low
Cappuccio et al[9]	Internal Medicine	√	√	pro low shift length and weekly working hours	Insignificant causality	High
Bilimoria et al[11]	Surgery	√		Not applicable	Insignificant causality	High
Landrigan et al[14]	Internal Medicine	√	√	pro low shift length and weekly working hours	Insignificant causality	High

Several studies have shown that long working hours or sleep deprivation deteriorated cognitive function of human brains.[2,4,14,15,19,20,23] In a survey from Mustahsan et al,[19] 20 % of the house officers and postgraduate trainees reflected that sleep deprivation was the risk factor of medical errors. On the other hand, dual positive results, improved sleep hygiene and less medical errors, were revealed after the implementation of the EWTD Rota in Cappuccio et al.[9]

Discontinued patient care is considered to be emerging risk factors of communicational medical errors following the diminishment of duty hours.[23] However, O'Leary et al[17] indicates otherwise: adverse events are not associated with the handover period because patients were taken care of by medical teams and all medical actions were under the surveillance of attending physicians. On the other hand, prolonged shifts or night shifts may decrease the cognitive function of residents or affects their circadian rhythm.[24,25] However, the question of the inconsistent result of the surgical department remains unsolved. Naylor et al[22] for example, demonstrated that most procedure-related complication after surgery occurred during the first 18 hours of their duty, attributing the insignificant difference between control and the intervention group to several factors such as the cavalier attitude, more team surveillance, and the different degrees of difficulty in a surgery. Since it is a sophisticated question, more studies are expected for a conclusion to be made.

Regarding this study, all sources were retrieved from the EMBASE and PubMed with access authorized by the National Taiwan University. All screening and selection processes were accomplished by the primary author (C W-Y), some articles prone to subjective filtration. The result of our study may not be generalized to other clinical specialties because of post-graduate trainees and residents are primary subjects of interests among the included studies, similarly inapplicable in countries other than the U.S. or the U.K.

Given the lack of eligible studies and the diverse background settings across studies, meta-analysis is not possible to be proceeded. Nevertheless, literature extraction in this study has provided a lens by systematic approach into either factor involved and into how duty hour restriction policy impacts patient safety. More researches are awaited for the evidence-based suggestion for the schedule change to reduce the fatigue or the workload of the medical staff as well as to benefit patients.

ACKNOWLEDGMENTS

The study stemmed from a research program of social medicine, which had been started up by Osaka University and National Taiwan University in October, 2017. Chao Wei-Yo (C W-Y): as the primary author, is responsible for the integrity of the article. Lee Chien-Cheng (L C-C), Chen Shyr-Chyr (C S-C) and Liprovided Cheng-Han (L C-H): substantial knowledge supports in biostatistics and epidemiology. Geerdes Emma (G E): revised the section of discussion of the manuscript. Yamazaki Mai (Y M), Tanoue Haruna (T H) and Hsiung Jo-Ching (H J-C): team members in the research program of social medicine, are appreciated for their contribution to the research and their support of the result from their research of "the association between mental depression and long working hours among physicians and nurses." Searching strategy for the PubMed and EMBASE was advised by Mrs. Hsin-Ping Chiu from the National Taiwan University Medical Library. All articles were obtained with full access authorized by the National Taiwan University Medical Library.

FUNDING

This work is sponsored by MOT research grant of Taiwan.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

1. Grober ED, Bohnen JM. Defining medical error. *Can J Surg.* 2005; 48(1): 39-44.

2. Arnedt JT, Owens J, Crouch M, Stahl J, Carskadon MA. Neurobehavioral performance of residents after heavy night call *vs* after alcohol ingestion. *JAMA.* 2005; 294(9): 1025-1033. doi: 10.1001/jama.294.9.1025

3. Lockley SW, Cronin JW, Evans EE, et al. Effect of reducing interns' weekly work hours on sleep and attentional failures. *N Engl J Med.* 2004; 351(18): 1829-1837. doi: 10.1056/NEJMoa041404

4. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Ann Intern Med.* 2009; 151(4): 264-269. doi: 10.1136/bmj.b2535

5. Critical Appraisal Skills Programme (CASP). 2015. Critical appraisal skills programme (CASP). Web site. http://www.casp-uk.net/. Accessed January 11, 2015.

6. Ouyang D, Chen JH, Krishnan G, Hom J, Witteles R, Chi J. Patient outcomes when housestaff exceed 80 hours per week. *Am J Med.* 2016; 129(9): 993-999. doi: 10.1016/j.amjmed.2016.03.023

7. Anderson JE, Goodman LF, Jensen GW, Salcedo ES, Galante JM. Restrictions on surgical resident shift length does not impact type of medical errors. *J Surg Res.* 2017; 212: 8-14. doi: 10.1016/j.jss.2016.12.038

8. Yaghoubian A, Kaji AH, Ishaque B, et al. Acute care surgery performed by sleep deprived residents: Are outcomes affected?. *J Surg Res.* 2010; 163(2): 192-196. doi: 10.1016/j.jss.2010.04.011

9. Bhavsar J, Montgomery D, Li J, et al. Impact of duty hours restrictions on quality of care and clinical outcomes. *Am J Med.* 2007; 120(11): 968-974. doi: 10.1016/j.amjmed.2007.07.026

10. Cappuccio FP, Bakewell A, Taggart FM, et al. Implementing a 48 h EWTD-compliant rota for junior doctors in the UK does not compromise patients' safety: Assessor-blind pilot comparison. *QJM.* 2009; 102(4): 271-282. doi: 10.1093/qjmed/hcp004

11. Landrigan CP, Rothschild JM, Cronin JW, et al. Effect of reducing interns' work hours on serious medical errors in intensive care units. *N Engl J Med.* 2004; 351(18): 1838-1848. doi: 10.1056/NEJMoa041406

12. Bilimoria KY, Chung JW, Hedges LV, et al. National cluster-randomized trial of duty hour flexibility in surgical training. *N Engl J Med.* 2016; 374(8): 713-727. doi: 10.1056/NEJMoa1515724

13. Fletcher KE, Davis SQ, Underwood W, Mangrulkar RS, McMahon LF, Saint S. Systematic review: Effects of resident work hours on patient safety. *Ann Intern Med.* 2004; 141(11): 851-857. doi: 10.7326/0003-4819-141-11-200412070-00009

14. Qureshi AU, Ali AS, Hafeez A, Ahmad TM. The effect of consecutive extended duty hours on the cognitive and behavioural performance of paediatric medicine residents. *J Pak Med Assoc.* 2010; 60(8): 644-649.

15. Gohar A, Adams A, Gertner E, et al. Working memory capacity is decreased in sleep-deprived internal medicine residents. *J Clin Sleep Med.* 2009; 5(3): 191-197.

16. Denson J, McCarty M, Fang Y, Uppal A, Evans L. Increased mortality rates during resident handoff periods and the effect of ACGME duty hour regulation. *Chest.* 2014; 146(4): 338A. doi: 10.1378/chest.2058935

17. O'Leary KJ, Turner J, Christensen N, Ma M, Lee J, Williams MV, Hansen LO. The effect of hospitalist discontinuity on adverse events. *J Hosp Med.* 2015; 10(3): 147-151. doi: 10.1002/jhm.2308

18. Gonzalo J, Kuperman E, Chuang CH, Lehman EB, Abendroth TW. Impact of an overnight internal medicine academic hospitalist program on patient outcomes. *J Gen Intern Med.* 2015; 30(12): 1795-1802. doi: 10.1007/s11606-015-3389-0

19. Mustahsan SM, Ali SM, Khalid F, et al. Sleep deprivation and its consequences on house officers and postgraduate trainees. *J Pak Med Assoc.* 2013; 63(4): 540-543.

20. Maltese F, Adda M, Bablon A, et al. Night shift decreases cognitive performance of ICU physicians. *Intensive Care Med.* 2016; 42(3): 393-400. doi: 10.1007/s00134-015-4115-4

21. Landrigan CP, Barger LK, Cade BE, Ayas NT, Czeisler CA. Interns' compliance with accreditation council for graduate medical education work-hour limits. *JAMA.* 2006; 296(9): 1063-1070. doi: 10.1001/jama.296.9.1063

22. Naylor RA, Rege RV, Valentine RJ. Do resident duty hour restrictions reduce technical complications of emergency laparoscopic cholecystectomy?. *J Am Coll Surg.* 2005; 201(5): 724-731. doi: 10.1016/j.jamcollsurg.2005.06.271

23. Gander P, Millar M, Webster C, Merry A. Sleep loss and performance of anaesthesia trainees and specialists. *Chronobiol Int.* 2008; 25(6): 1077-1091. doi: 10.1080/07420520802551428

24. Amirian I, Andersen LT, Rosenberg J, Gogenur I. Working night shifts affects surgeons' biological rhythm. *Am J Surg.* 2015; 210(2): 389-395. doi: 10.1016/j.amjsurg.2014.09.035

25. Amirian I, Andersen LT, Rosenberg J, Gogenur I. Laparoscopic skills and cognitive function are not affected in surgeons during a night shift. *J Surg Educ.* 2014; 71(4): 543-550. doi: 10.1016/j.jsurg.2013.12.007

APPENDIX

• Searching Strategy for Embase

('health care error'/exp OR 'health care error' OR 'healthcare error'/exp OR 'healthcare error' OR 'medical error'/exp OR 'medical error' OR 'medical errors'/exp OR 'medical errors' OR 'adverse effect'/exp OR 'adverse effect' OR 'adverse event'/exp OR 'adverse event' OR 'adverse events'/exp OR 'adverse events' OR 'adverse reaction'/exp OR 'adverse reaction' OR 'serious medical error' OR 'patient safety'/exp OR 'patient safety') AND ('interns and residents'/exp OR 'interns and residents' OR 'resident'/exp OR 'resident' OR 'resident doctor'/exp OR 'resident doctor' OR 'resident physician'/exp OR 'resident physician' OR 'resident surgeon'/exp OR 'resident surgeon' OR 'surgery resident'/exp OR 'surgery resident' OR 'surgical resident'/exp OR 'surgical resident' OR 'doctor'/exp OR 'doctor' OR 'medical doctor'/exp OR 'medical doctor' OR 'medical practitioner'/exp OR 'medical practitioner' OR 'physician'/exp OR 'physician' OR 'physician associate'/exp OR 'physician associate' OR 'physicians'/exp OR 'physicians' OR 'practitioner'/exp OR 'practitioner' OR 'private physician'/exp OR 'private physician') AND ('long working hour' OR 'duty hours' OR 'accreditation council for graduate medical education' OR 'duty hour rule' OR 'duty-hour limitation')

• Searching Strategy for PubMed

("Medical Errors"[Mesh] OR "adverse events" OR "adverse effects" OR "serious complications" OR "serious events" OR "postoperative complications") AND ("long working hours" OR "duty hours" OR "Duty hour rule" OR "ACGME") AND ("Physicians"[Mesh] OR "Physicians, Primary Care"[Mesh] OR "Osteopathic Physicians"[Mesh] OR "Occupational Health Physicians"[Mesh] OR "Physicians, Women"[Mesh] OR "Physicians, Family"[Mesh] OR "Practice Patterns, Physicians"[Mesh] OR "Physician's Role"[Mesh] OR "Medical Staff, Hospital"[Mesh] OR "General Practitioners"[Mesh] OR "Internship and Residency"[Mesh])