

What Is Known: Examining the Empirical Literature in Resident Work Hours Using 30 Influential Articles

Ingrid Philibert, PhD, MBA

ABSTRACT

Background Examining influential, highly cited articles can show the advancement of knowledge about the effect of resident physicians' long work hours, as well as the benefits and drawbacks of work hour limits.

Objective A narrative review of 30 articles, selected for their contribution to the literature, explored outcomes of interest in the research on work hours—including patient safety, learning, and resident well-being.

Methods Articles were selected from a comprehensive review. Citation volume, quality, and contribution to the evolving thinking on work hours and to the Accreditation Council for Graduate Medical Education standards were assessed.

Results Duty hour limits are supported by the scientific literature, particularly limits on weekly hours and reducing the frequency of overnight call. The literature shows declining hours and call frequency over 4 decades of study, although the impact on patient safety, learning, and resident well-being is not clear. The review highlighted limitations of the scientific literature on resident hours, including small samples and reduced generalizability for intervention studies, and the inability to rule out confounders in large studies using administrative data. Key areas remain underinvestigated, and accepted methodology is challenged when assessing the impact of interventions on the multiple outcomes of interest.

Conclusions The influential literature, while showing the beneficial effect of work hour limits, does not answer all questions of interest in determining optimal limits on resident hours. Future research should use methods that permit a broader, collective examination of the multiple, often competing attributes of the learning environment that collectively promote patient safety and resident learning and well-being.

Editor's Note: The ACGME News and Views section of JGME includes data reports, updates, and perspectives from the ACGME and its review committees. The decision to publish the article is made by the ACGME.

Introduction

Resident work hours have been considered important to the learning and professional socialization of physicians, in an industry where patient needs are not met during the standard workday. Yet these long hours also are of concern for patients and residents, due to the association between sleep loss, medical errors for patients, and risks to residents themselves, including motor vehicle accidents, occupational injuries, and negative effects on well-being. Ideally, decisions regarding work hour limits for resident physicians should be based on the relevant scientific literature regarding approaches that optimize patient safety, resident learning, and resident safety and well-being.

The Accreditation Council for Graduate Medical Education (ACGME) work hour standards, which include the early standards that predate the 2003

common program requirements and the standards that went into effect in 2011, were guided by the available scientific evidence generally gleaned from published studies.^{1,2} The aim of this selective, narrative review is to present the seminal articles over the past 4 decades that shaped the community's and the ACGME's thinking about resident hours, including highlighting themes and controversies in this literature. An additional aim is to identify areas where relevant literature is lacking, and suggest directions for future research.

Methods

The review encompasses 30 well-cited, influential articles published in US peer-reviewed journals between 1971 and 2013. Articles were selected from a comprehensive review of the literature on resident work hours and related dimensions. These were compiled by the author for an ACGME Task Force that was charged with a revision of Section VI (Resident Duty Hours in the Learning and Working Environment) of the Common Program Requirements. This compilation consisted of more than 1050 original articles. To be included in this narrative review, a study had to examine resident hours or limits on them, and had to be a frequently

DOI: <http://dx.doi.org/10.4300/JGME-D-16-00642.1>

cited article that addressed an aspect of resident work hours, such as medical errors, resident safety and well-being, and resident learning, or the positive or negative consequences of work hour limits. Articles also had to meet a basic quality criterion.

Citation volume was assessed independently using Google Scholar. For quantitative studies, quality was examined using the Medical Education Research Study Quality Instrument (MERSQI).³ The MERSQI assesses design, sampling, type of data, validity, data analysis, and outcomes, for a maximum score of 18 and an average score of 9.95 for medical education studies. To be included, studies had to score 11 or above, and the scores for the 28 quantitative articles ranged from 11.4 to 17, with many of the older studies scoring lower on the metric. MERSQI data are not reported individually, and a number of studies have separate scores for different outcome variables, such as the external collection of safety data versus the self-reporting of work hours or educational outcomes. The 2 purely qualitative studies in the analysis scored high on a quality assessment tool for qualitative research.⁴ The organization of the review highlights the historical evolution of the community's knowledge and understanding about resident hours, the focus on several constant themes throughout the 4 decades, and the emerging focus on new outcomes of interest.

Results

As shown in the TABLE, many of the articles included in this review are among the most highly cited articles in this vast literature, with the most cited study garnering nearly 1500 citations.^{4-11,13-30,32-34} In contrast, 2 important qualitative studies included in the analysis^{12,31} received fewer citations.

The major themes in the early literature on work hours begin the discussion of the impact of long hours on residents' clinical and cognitive performance,^{4,8} resident safety and well-being,^{5,17,18,23,24,26} and resident driving safety after call or night shifts.^{14,16,21} A second theme encompasses the risks to patients associated with resident work,^{10,20,22,32-24} including higher risks for night admissions⁹ and problems with cross-coverage of patients.¹³ In contrast, data on the impact of duty hour limits on resident learning and professional development are scarce, and studies generally are limited to assessing the impact on board and in-training examination performance and resident self-reporting of educational impact.^{8,29,30}

In 1989, New York became the first state to regulate resident hours; thus, the third theme is the impact of limits on resident hours on the quality and safety of care. Studies of the impact of New

York State's regulation of resident hours predate studies of limits at the national level. These early studies produced some mixed results, including delayed test ordering and increased complications in medical residents' patients,¹⁰ as well as residents reporting conflict between their desire to comply with the regulations and their obligation to patients and colleagues.¹² While these early studies clearly highlight the need for limits on resident work, beginning with that by Friedman and colleagues,⁴ they also present conflicting findings, with no apparent impact of long hours for surgical residents⁷ and an increased number of errors for surgical patients under the New York State limits.²² This marked the beginning of a 3-decades-long debate about differences in the impact of work hour limits in medical versus surgical specialties. Studies of the effect of 2003 ACGME standards generally suggest a benefit of these limits, particularly the 80-hour weekly limits and standards reducing the frequency of overnight call. However, some studies produced mixed findings or suggested no impact on quality and safety of care. For example, 2 large studies using administrative data showed no significant differences in outcomes for medical²⁸ or surgical^{26,28} patients.

Several studies showed that residents' actual hours that were worked routinely exceeded scheduled hours, as residents tend to remain to complete tasks. A high-quality study of internal medicine interns limited to 16 hours in an intensive care unit rotation showed that actual hours worked significantly exceeded scheduled hours during the intervention and control weeks¹⁹; 2 studies showed that residents regularly remained beyond scheduled hours to complete clinical tasks^{12,31}; and an early study showed significant differences in resident work styles and efficiency, with longer and more difficult on-call experiences for some, despite comparable objective workloads.¹¹

Data on resident well-being indicate serious consequences of residency to trainees' well-being, and suggest that the intern year is a particularly stressful time.^{6,18,24} Studies also show a benefit of the 2003 ACGME standards on well-being in both medical and surgical residents.^{23,26} Research on the 2003 and 2011 limits clearly show that internal medicine residents' weekly hours have declined from more than 85 at the implementation of the 2003 ACGME standards to around 65 hours per week. This reduction in weekly duty hours may be most responsible for the benefits attributed to the 2003 and 2011 common requirements, offering empirical support for the 80-hour weekly limit and limits on the frequency of overnight call. In this context, it is

interesting to note that 6 of 22 primary studies included in a systematic review of the benefit of a 16-hour limit³⁵ actually assessed the impact of reductions in call frequency (while retaining 24-hour call), with call patterns and associated hours in the intervention groups approximating the current practice of call every third or fourth night. A concerning finding is that this reduction in weekly hours, and compliance with the work hour limits, appears to have resulted in an increase in clinical demands, increased perceived workload and stress, and reduced participation in educational activities,^{32–34} with 1 study finding an increase in self-reported errors by first-year residents.³⁴

The literature clearly shows the negative effect of sleep loss: several highlight that in a 24-hour/7-day industry, under almost any model of work scheduling, including shift-based approaches, some individuals must work at their circadian nadir, with associated consequences for their performance, safety, and well-being.^{16,17,20,24} Scheduling may mitigate this decrement, but it is not possible to completely eliminate it. Protected time for napping has been suggested as an intervention; however, 2 studies showed that efforts to “protect” sleep time for on call did not increase sleep, as residents may not use the nap option out of a sense of professionalism and a desire to provide care for their patients.^{15,25}

Studies that offer insight into the educational impact of work hour limits are scant, with 1 study finding a reduction in performance on the certifying examination for neurological surgery residents,³⁰ and several finding reductions in residents’ self-reported participation in didactic activities and satisfaction with their education.^{33,34}

Discussion

The utility of this influential literature is limited due to small sample sizes, the short-term nature of many interventions, as well as concerns about the ability to generalize from studies performed in a single specialty and, often, a very specific clinical context, such as an intensive care unit, to all specialties and clinical contexts. In fact, 17 of the 30 influential studies used internal medicine residents, and 12 used solely first-year internal medicine trainees, potentially echoing a critique of the psychology literature and its use of sophomore-level college student subjects.³⁶ Three studies each involved residents across multiple specialties, 3 involved surgery residents, 2 involved medicine and surgery residents, 2 involved pediatrics residents, and 1 each involved residents in emergency medicine, family medicine, and neurological surgery.

Many medical, surgical, and hospital-based specialties are represented in a limited way in the large data set from which these studies were selected. In addition, to date few studies have assessed the large number of additional standards added in 2011, including enhanced standards for supervision, the creation of a safe and professional learning environment, educating residents and faculty on fatigue and alertness management, and resident participation in quality and safety improvement.

It also is important to note that positive or negative changes in any outcomes of interest cannot be unequivocally attributed to the limits. Concurrent with the implementation of the 2 sets of common standards, there has been a substantial increase in a focus on quality and safety in teaching settings, as part of the standards themselves,^{1,2} through the ACGME’s Clinical Learning Environment Review,³⁷ and through a multitude of efforts within the health care sector to enhance the safety and quality of care.³⁸ There also is some indication that the improvement or worsening of quality and safety indicators may be related to greater resident peer or faculty involvement in care. This may enhance patient safety, but it could be problematic for learning. For real learning in a clinical context, residents need to be immersed in systems of patient care, not bobbing gently at the surface in the interest of safety. A letter to the editor on a widely cited trial²⁰ of a 16-hour limit on continuous duty attributed the beneficial outcomes to increased involvement and coverage of first-year residents’ work by more senior residents and faculty.³⁹

Efforts to address the negative impact of shortened work hours, particularly the 16-hour limit for first-year residents instituted in 2011, have included night float, improved handoff procedures, and teaching on the night shift, among others. None were universally effective in addressing the negative effects of the work hour limits on patient care or learning.^{13,25,32–24} In fact, some studies have found that night float may produce outcomes inferior to those in other forms of scheduling, and in 1 study night float was terminated early due to concerns about quality and safety of care.³³

To deny the negative impact of long work hours, sleep loss, and fatigue in residents is folly, as these physiological demands are very real. At the same time, to have these physiological limits become the major driver of the system to ensure the learning, professional development, and socialization of the next generation of physicians likely is unrealistic as well. While empirical data is scarce, there is a pervasive sense that the ACGME common work

TABLE
Study Findings and Contributions to/Impact of the Knowledge About Work Hour Limits

Source, y	Findings/Contributions	Study Attributes	Observations
1. Friedman et al, ⁴ 1971 (cited by 345)	<i>First study to show sleep loss affects clinical performance postcall in first-year internal medicine residents. Postcall first-year residents performed less well on reading a standardized electrocardiogram compared to rested colleagues, and exhibited negative mood.</i>	14 first-year internal medicine residents; residents not on call the night before testing served as controls.	Every other night call with high potential for chronic sleep loss in subjects and controls. In 1989–1990, ACGME instituted standards limiting call to every third night.
2. Reuben, ⁵ 1985 (cited by 165)	<i>First to quantify depressive symptoms and link it to residency year and training context. Study showed higher prevalence of depressive symptoms in first-year residents (28.7% compared to 21.4% overall), falling with each year of training.</i>	68 internal medicine residents across all 3 y; monthly self-reports for 1 y (87.3% response rate).	Shows clinical context affects depressive symptoms, with a prevalence of 34.8% for first-year residents on wards, and 33% for residents on intensive care rotations.
3. Deaconson et al, ⁶ 1988 (cited by 169)	<i>First to show that there is no significant effect of postcall status on the performance of surgery residents. Postcall surgery residents who served as their own controls showed no significant reduction in performance on standardized psychometric and visual-haptic tests.</i>	26 surgical residents in 3 cohorts, serving as their own controls, studied on postcall and alternate days.	Every other night call with high potential for chronic sleep loss, with less than 4 h of sleep during the preceding 24 h in 89% of postcall subjects.
4. Jacques et al, ⁸ 1990 (cited by 103)	<i>First to show sleep loss affects performance on a standardized examination. Family medicine residents showed a significant decline in in-training examination scores with decreasing sleep the night before, with effect of a loss of 1 night sleep similar to the difference in scores between first- and third-year residents.</i>	353 family practice residents in 21 programs; residents call schedules and prior sleep varied in the sample.	Authors suggested that testing over several hours may be necessary to detect subtle but significant differences in cognitive performance with relatively mild degrees of sleep loss.
5. Hillson et al, ⁹ 1992 (cited by 53)	<i>First to show that evening/night admissions to teaching units had increased relative risk inpatient mortality (1.21; P = .03), decreased hospital length of stay (8.1%; P < .0001), and increased total charges (3.1%; P = .007).</i>	22 112 patients on the internal medicine service of a teaching hospital admitted by an on-call first-year resident between January 1, 1980, and December 31, 1987.	No linear relation was found between number of admissions and length of stay. Analysis of nonlinear effects showed length of stay rose, then fell as first-year residents received more on-call admissions.
6. Laine et al, ¹⁰ 1993 (cited by 247)	<i>First to show that work hour limits in New York State were associated with delayed test ordering by internal medicine residents and increased in-hospital complications. There were no statistically significant differences in more serious outcomes.</i>	263 patients discharged from a New York hospital in October 1988, and 263 discharged from the same service in October 1989.	Suggested a need to further determine the effects of staffing changes to optimize quality of patient care and resident education.
7. Tanz and Charrrow, ¹¹ 1993 (cited by 31)	<i>First to show an association between working style, hours of sleep, and self-perceived workload for residents, with some residents consistently reporting they worked more than colleagues, and having a reputation among colleagues for having difficult on-call experiences. A reputation for difficult on call was associated with less sleep (r = -0.77) but no difference in objective workload.</i>	Self-reports on 1355 on-call experiences of 19 first-year pediatrics residents during the 1984–1985 academic year.	The absence of an association between perceived and actual workload is attributed to large differences in the residents' efficiency and working styles.
8. Yedidia et al, ¹² 1993 (cited by 37)	<i>First to show that first-year residents under work hour limits have competing priorities in complying with the limits and providing care to patients. Interns were concerned about leaving patients at critical times, confidence in colleagues, concerns for colleagues' workload, and concern for educational impact of reduced hours.</i>	Interviews with 21 of 24 first-year internal medicine residents on an inpatient rotation in a New York urban teaching hospital.	Suggested a conflict between the values of autonomy and individual accountability inculcated in residents and the collective responsibility required under work hour limits.
9. Petersen et al, ¹³ 1994 (cited by 629)	<i>First to show that cross coverage of inpatient units to reduce resident staffing and hours on an internal medicine resident service at night was associated with a higher likelihood of an adverse event. In 26% of potentially preventable adverse events, patients were covered by a physician from another team (compared with 12% for coverage by a familiar resident [OR = 3.5; P = .01]).</i>	3146 patients admitted over a 4-mo period, self-reporting of adverse events, board-certified internists confirmed events. Of 124 adverse events, 44% were judged to be preventable.	Highlighted concerns with scheduling patterns that reduce continuity and perceptions of patient ownership, including night float, which had become common as a strategy to limit call and a recruitment advantage for internal medicine programs.

TABLE
Study Findings and Contributions to/Impact of the Knowledge About Work Hour Limits (continued)

Source, y	Findings/Contributions	Study Attributes	Observations
10. Marcus and Loughlin, ¹⁴ 1996 (cited by 190)	<i>First to state that postcall residents are significantly more likely to fall asleep while driving, and are more likely to be involved in motor vehicle accidents, and 49% of residents surveyed had fallen asleep at the wheel, with 90% of events occurring postcall.</i>	70 pediatric residents on call every fourth night, and 85 faculty members, rarely disturbed at night.	The debate about postcall driving safety for residents continued, and became a requirement for sleep facilities and/or safe transportation options for fatigued residents in the 2011 common program requirements.
11. Richardson et al, ¹⁵ 1996 (cited by 92)	<i>Showed an intervention to "protect" sleep time on call does not increase sleep or enhance alertness in first-year residents. Interns in the protected sleep group and in the control group spent on average less than 5 h in bed trying to sleep, and slept a mean of 3.67 h. Measures of alertness and performance did not differ significantly between the 2 groups.</i>	26 first-year internal medicine residents assessed throughout a 36-h call day. Residents were provided with 4 h of protected time for sleep by a covering night float resident.	Authors highlighted chronic sleep deprivation in the intervention group as a possible reason for the lack of effect of a protected nap period. A statement strongly encouraging strategic napping was included in the 2011 ACGME standards.
12. Steele et al, ¹⁶ 1999 (cited by 175)	<i>Showed that residents in a specialty with regulated shift length still are at risk for motor vehicle accidents driving home after a shift. For second- to fourth-year emergency medicine, nearly three-fourths of the motor vehicle accidents and 80% of the near-crashes followed the night shift.</i>	1554 residents in 78 emergency medicine programs, with 76 residents (8%) reporting 96 crashes, and 553 (58%) reporting 1446 near-crashes.	Shows that driving home after night work is an occupational risk for residents on shift-based schedules. The study also is the first to show trait in the response to sleep deprivation and night shift work.
13. Parks et al, ¹⁷ 2000 (cited by 72)	<i>Showed a 1.5-fold higher risk of sustaining a blood-borne pathogen exposure (needle stick, laceration, or splash) for resident physicians and medical students working during night hours, compared with daytime hours. Average rate of exposures was 40 per hour per 1000 doctors in training during the day, and 60 incidents per hour between 6:00 PM and 5:59 AM.</i>	Retrospective review of 745 reported blood-borne pathogen exposures in a large urban teaching institution between November 1993 and July 1998.	Day- and nighttime patterns in exposure are consistent with studies of exposure in other health professions working shorter shifts.
14. Bellini et al, ¹⁸ 2002 (cited by 337)	<i>First to quantify the reduction in enthusiasm and empathy and an increase in personal distress for first-year internal medicine residents from entry into residency. Residents showed less tension, depression, anger, and fatigue, and more vigor, than the general population, better perspective taking ($P < .001$) and empathic concern, and lower personal distress. Five months into training, scores showed increases in the depression-dejection ($P < .001$), anger-hostility ($P < .001$), and fatigue-inertia ($P < .001$), a decrease in empathy ($P = .005$), and an increase in distress ($P < .001$).</i>	61 first-year internal medicine residents at 1 university-based medical center, sampled in June, November, and February for Profile of Mood States and Interpersonal Reactivity Index.	This is one of a small number of high-quality studies to measure the reduction in altruism and enthusiasm after the transition from medical school to residency.
15. Lockley et al, ¹⁹ 2004 (cited by 802)	<i>Showed that during a crossover trial of a 16-h limit on continuous work, 17 of the 20 first-year residents worked significantly more than 80 h/wk in the control (traditional schedule) portion of the study (mean hours, 84.9; range, 74.2–92.1), and all worked significantly less than 80 h in the intervention portion (mean, 65.4; range, 57.6–76.3). During the intervention residents had less than half the attentional failures while at night ($P = .02$) than in the traditional schedule.</i>	20 first-year internal medicine residents on two 3-wk intensive care unit rotations, serving as their own controls.	It is not clear whether the 16-h schedule or the significant reduction in weekly work hours by nearly 20 h and a 5.8-h increase in weekly sleep are responsible for the outcomes. For both the intervention and control periods work hours significantly exceeded "scheduled" hours.
16. Landrigan et al, ²⁰ 2004 (cited by 1451)	<i>First to quantify the errors made by first-year residents under a 16-h limit on continuous work hours and those made by first-year residents working a traditional overnight call schedule. Interns made 36% fewer serious medical errors during the intervention schedule (136.0 versus 100.1 per 1000 patient-days), including 57% fewer nonintercepted serious errors.</i>	Direct observation and data collection for 2203 patient-days involving 634 admissions	In this very high-quality study, it is not possible to conclusively attribute the reduction in errors to the elimination of > 24 h call versus the significant reduction in weekly hours and increase in sleep for residents in the intervention group.
17. Barger et al, ²¹ 2005 (cited by 742)	<i>First large, multispecialty study to show that odds for reporting a motor vehicle crash or a near-miss incident were 2.3 (95% CI 1.6–3.3) and 5.9 (95% CI 5.4–6.3), respectively, after an extended work shift. In months in which interns worked 5 or more extended shifts, the risk of falling asleep while driving was significantly increased (OR = 2.39 [95% CI 2.31–2.46] and 3.69 [95% CI 3.60–3.77], respectively).</i>	National survey of 2737 first-year residents who completed 17003 monthly reports that provided information about work hours and shifts, and incidents of sleeping while driving, motor vehicle crashes, and near-misses.	This finding contributed to the addition of a requirement for sleep facilities and/or safe transportation options for fatigued residents in the 2011 ACGME common program requirements.

TABLE
Study Findings and Contributions to/Impact of the Knowledge About Work Hour Limits (continued)

Source, y	Findings/Contributions	Study Attributes	Observations
18. Poulouse et al, ²² 2005 (cited by 114)	<i>First study to show a negative impact of resident work hour limits on standardized patient safety indicators for surgical patients. Increased rates for accidental puncture or laceration (0.15; 95% CI 0.09–0.20; $P < .05$) and pulmonary embolus or deep venous thrombosis (0.43; 95% CI 0.03–0.83; $P < .05$) were observed after institution of work hour limits, with no change in control institutions. Analysis encompassed a range of post-operative complications.</i>	Interrupted time series analysis of an average of 2.6 million discharges per year from New York teaching hospitals between 1995 and 2001, compared with New York nonteaching and California teaching hospitals.	Contributed to an emerging debate over differences in the impact of work hour limits on surgical versus medical specialties, with surgical programs experiencing a negative effect under the limits.
19. Gopal et al, ²³ 2005 (cited by 214)	<i>First to show a reduced burnout for internal medicine residents under the 2003 work hour standards. As self-reported hours for residents decreased from a mean of 74.6 to 67.1, 13% fewer residents experienced high emotional exhaustion (42% versus 29%), and a trend toward fewer residents with high depersonalization (61% versus 55%) or a positive depression screen (51% versus 41%).</i>	Survey of all internal medicine residents at a large academic medical center in May 2003 (121 of 139 responded) and May 2004 (106 of 143 responded).	Residents reported attending fewer educational conferences per month (18.99 versus 15.56, $P = .01$), and overall residency satisfaction decreased 6 mm on a 100-mm visual analogue score ($P = .02$).
20. Rosen et al, ²⁴ 2006 (cited by 249)	<i>First to detect an association between becoming chronically sleep deprived and becoming depressed (OR = 7; $P = .014$). Study compared the beginning and end of internship, prevalence of "high" chronic sleep deprivation scores increased (9% to 43%; $P = .0001$), prevalence of moderate depression increased (4.3% to 29.8%; $P = .0002$), and the percentage of residents reporting a high level of burnout increased (4.3% to 55.3%; $P < .0001$). Scores for empathy originally were more favorable but approached norms at the end of the year ($P = .15$).</i>	In July 2002, 47 first-year internal medicine residents completed the sleepiness, depression, and burnout surveys, and completed the same instrument at the end of internship in June 2003.	Study predated the July 1, 2003, implementation of common work hour standards, but internal medicine operated under an 80-h weekly limit. Suggested a small percentage of trainees enter residency with some degree of burnout and/or depression.
21. Arora et al, ²⁵ 2006 (cited by 129)	<i>Showed that, while first-year internal medicine residents on call with an on-duty nap intervention received an average of 41 min additional in-hospital sleep compared with residents on a standard call schedule, many residents on the nap schedule opted not to use coverage, citing their desire to care for their patients and concerns about continuity of care.</i>	From July 2003 to June 2004, residents were assigned for 2 wk monthly to a nap schedule that provided coverage from 12 AM to 7 AM; the other 2 wk were a standard call schedule.	Suggests socialization to patient responsibilities as a barrier to residents accepting a protected nap and coverage option during on-call nights.
22. Hutter et al, ²⁶ 2006 (cited by 286)	<i>First to show that the 2003 work hour limits increased self-reported sleep, enhanced motivation, and reduced burnout in surgical residents, with significantly less "emotional exhaustion," changing from 29.1 "high" to 23.1 "medium"; $P = .02$. There was no difference in the quality of patient care (NSQIP data), no reduction in operative volume or ABSITE scores.</i>	Four prospective analyses at 1 academic surgical program, with data collected both before and after the institution of the ACGME work hour limits.	<i>Highlighted the ability of NSQIP data to assess in surgical outcomes, creating a "midlevel" metric between single-site studies and large sample studies using administrative data.</i>
23. Singh et al, ²⁷ 2007 (cited by 301)	<i>Showed medical errors involving trainees are more likely to relate to lack of supervision and teamwork and handoff problems than to sleep loss or fatigue. Most errors involving trainees related to judgment (72%), teamwork (70%), and clinical/technical competence (58%). Fatigue, excessive workload, and inadequate staffing still were disproportionately present in errors involving trainees (5% versus 1% and 19% versus 5%, respectively).</i>	Analysis of 240 closed malpractice claims from 5 insurers that involved trainees.	Suggests a broad host of factors in trainee errors, with trainee errors having more contributing factors than those not involving trainees (mean of 3.8 versus 2.5; $P < .001$).
24. Volpp et al, ²⁸ 2007 (cited by 279)	<i>Showed the 2003 ACGME work hour limits were not associated with significant worsening or improvement in mortality for Medicare patients in the first 2 y after their implementation (academic years 2003–2005) to the 3 y prior to the work hour limits (academic years 2000–2003). Compared with nonteaching hospitals, teaching-intensive hospitals had 4.4% relative increase in mortality over the prior period for medical patients and a 2.3% relative increase for surgical patients, with neither being statistically significant.</i>	Interrupted time series analysis of 8.5 million Medicare patients with common admission diagnoses, or general, orthopedic, or vascular surgery procedures admitted to 3321 US nonfederal hospitals between 2000 and 2005.	Showed that the ACGME work hour standards did not result in a significant improvement or worsening in outcomes for Medicare patients in the first 2 y after they were implemented.

TABLE

Study Findings and Contributions to/Impact of the Knowledge About Work Hour Limits (continued)

Source, y	Findings/Contributions	Study Attributes	Observations
25. Arora et al, ²⁹ 2008 (cited by 95)	<i>Showed an inverse relationship between workload, resident efficiency, sleep obtained, and participation in educational activities by first-year internal medicine residents during call, and suggested call early in the academic year as a particularly taxing time, with each new on-call admission associated with 10.5 min less sleep and 13.2 min added to shift duration. A higher number of patients on the service was associated with a lower odds of participation in educational activities. Weeknight call and call early in the academic year were associated with the most sleep loss and longest shift durations.</i>	Prospective cohort study of 56 first-year internal medicine residents at 1 institution studied from July 2003 to June 2005 (1100 call nights). Mean (SD) sleep duration on call was 2.8 (1.5) h, and mean (SD) shift duration was 29.9 (1.7) h. Interns spent 11% of time in educational activities.	Presents longitudinal quantitative information confirming the existence of a relationships between workload, sleep, and resident "efficiency" highlighted by Tanz and Charrow. ¹¹
26. Jagannathan et al, ³⁰ 2009 (cited by 107)	<i>Showed a negative effect of work hour limits on performance on the ABNS written examination. For residents taking the examination for self-assessment, performance decreased from 310 in 2002 to 259 in 2006 (16% decrease; $P < .05$), despite stable mean USMLE Step 1 scores. Despite an increase in the number of resident registrations to national meetings, the number of resident abstracts decreased from 345 in 2002 to 318 in 2007 (7% decrease; $P < .05$).</i>	ABNS and USMLE scores for all second- and third-year residents in ACGME-accredited neurological surgery training programs taking the examination between 2002 and 2006; data on meeting registration and abstracts presented by residents.	Highest-quality study of the impact of work hour limits on performance on a standardized examination.
27. Szymczak et al, ³¹ 2010 (cited by 43)	<i>Highlighted factors operating in resident decisions with regard to when to leave or when to stay, and residents' rationale for misreporting hours due to a perception that accreditors do not understand these nuances. Residents make careful, reasoned judgments about when to remain beyond the limits, are thoughtful about tradeoffs inherent in the regulations, and provide nuanced and complex reasons for misreporting their work hours.</i>	Ethnographic observation of internal medicine and general surgery residents in 2008, and in-depth interviews with 17 residents.	Suggests that concerns about residents exhibiting "shift work" mentality and lack of professionalism demonstrated by "falsifying" work hour reports, may be overstated.
28. McCoy et al, ³² 2011 (cited by 37)	<i>Showed that a 16-h limit on continuous work for second-year internal medicine residents left them feeling less prepared to manage cross coverage of patients ($P = .006$). Residents worked fewer weekly hours (64.3 versus 68.9 h; $P = .40$), but reported fewer episodes of 10 h off between shifts (24 versus 2 episodes; $P = .004$). There were no changes in mortality, readmissions, or length of stay, but there was a trend toward decreased perception of quality of education and balance of personal and professional life.</i>	Observational study of a staffing-neutral intervention that limited shifts to 16 h on 2 gastroenterology services over a 3-mo period. Patient care variables were collected from medical records; resident variables were collected from weekly surveys and end-of-rotation evaluations.	A staffing-neutral 16-h limit on continuous work hours may require some cross coverage of patients and may not be perceived well by residents, despite a reduction in weekly hours.
29. Desai et al, ³³ 2013 (cited by 119)	<i>Showed that 2 scheduling models compliant with the 2011 ACGME common standards increased sleep in first-year internal medicine residents, but reduced participation in didactics and availability during daytime hours, while increasing handovers. Increase in sleep was 5.1 versus 8.3 h for night float, and 7.5 versus 10.2 h for the call every fifth night model.</i>	3-mo crossover study of 43 first-year internal medicine residents randomly assigned to teams using either a 2003-compliant schedule or 2 models of a 2011-compliant schedule.	Showed significant problems with quality of care in the night float model, to such an extent that it was terminated early.
30. Sen et al, ³⁴ 2013 (cited by 91)	<i>Showed that first-year internal medicine residents, while working fewer hours under the 2011 ACGME standards, did not report more sleep or improved well-being, but reported an increase in self-reported medical errors (from 19.9% to 23.3%; $P = .007$). Self-reported weekly hours decreased from an average of 67.0 to 64.3, and hours slept decreased from 7.0 to 6.8; depressive symptoms and well-being showed no change.</i>	Longitudinal cohort study of 2323 interns in 51 programs, who self-reported work hours, hours of sleep, depressive symptoms, well-being, and medical errors at 3, 6, 9, and 12 mo, comparing 2009 and 2010 to 2011.	Assessment of only 1 y under the new standards makes it difficult to attribute the findings to a negative impact of the standards versus issues with systems adjusting to a new set of work hour limits.

Abbreviations: OR, odds ratio; ACGME, Accreditation Council for Graduate Medical Education; CI, confidence interval; NSQIP, National Surgical Quality Improvement; ABSITE, American Board of Surgery In-Training Examination; ABNS, American Board of Neurological Surgery; USMLE, US Medical Licensing Examination. NS, not significant.

hour standards, particularly those implemented in 2011, have resulted in “work compression,” by wringing out lower intensity work and downtime hours at the hospital that were valuable from an educational, professional socialization, and camaraderie perspective. This is not an indictment of the standards, but more likely of how they were implemented in some resource-constrained, resident-dependent hospitals. There also is some suggestion from citation patterns that certain areas initially thought promising have largely been abandoned as areas for study. For example, a high-quality study of a budget-neutral napping intervention, published in a leading medical education journal, received only 3 citations,⁴⁰ potentially because research priorities have shifted to other interventions with more community buy-in. The work hour reductions instituted in 2003 and 2011 should have brought about changes in the structure of residency training, but this does not appear to have happened in all programs and institutions.

New common work hour standards proposed by the ACGME on November 4, 2016,⁴¹ affirmed continued support for an 80-hour weekly limit, 1 day in 7 free of all program duties, and overnight call no more frequently than every third night. The new standards propose the elimination of the 16-hour limit on continuous duty. The decision was based on evidence that a universal, cross-specialty, cross-clinical setting application of a 16-hour limit has not been beneficial. It is important to note that programs and specialties that have found this limit effective, often as “ramp-up” to allow individuals to transition from the relatively limited clinical experience of many medical students, to the 24-hour, 7-day responsibilities of residency and medical practice, will be able to continue to use it. The new standards seek to address work compression and the shifting of clinical and administrative work to non-counted hours, as well as significantly increase the focus on resident and faculty well-being, and finding meaning in work.⁴¹

Limitations of this review include the potentially subjective selection of articles by the author, with different individuals curating this literature potentially reaching different conclusions as to what constitutes the most influential studies. While citation volume is a fairly undisputed metric, there are other highly cited articles, and there are no accepted metrics of the practical impact of research. In addition, the MERSQI assesses internal validity, yet its dimensions do not address generalizability. Several of the highest-quality studies were conducted with very small samples, in a single specialty, and often in a specific clinical context, reducing general-

izability. The review does not include studies on the impact of work hour limits on operative volume, due to the limitations of this literature and the fact that operative volume is a poor proxy for the quality of surgical education. Finally, a number of highly cited studies were omitted from the review because they used identical methods for slightly different populations (eg, patients in Veterans Affairs hospitals versus Medicare patients in nonfederal hospitals) or providing different publications of analyses of a larger data set.

There are a number of underinvestigated areas, with the most prominent being the impact of work hour limits on resident education. This likely is due to many potential confounding variables, and a dearth of accepted, useful metrics for this outcome. A requirement for safe practice in the future is a competent graduate, and a concern is that the safety of patients receiving care in teaching settings today might be achieved at the expense of the future patients that residents will care for after completing training. This will require more research into evidence-based ways to balance these competing demands. There also is a need for research that addresses continuity in residents' exposure to patients, to allow them to develop an understanding of the progression of disease, a patient's response to treatment, or the diagnosis and management of complications. A second area for future research entails addressing the effect of duty hour limits for surgical programs, as some studies suggest a negative impact in the areas of education and quality of patient care,⁴² and recent research suggests benefit in reverting to slightly more liberal standards.⁴³ Finally, studies often use historical comparisons that assume that other attributes of the learning environment have remained stable. However, patient acuity, intensity of service, new technology and, particularly, the widespread use of the electronic health record and resulting documentation needs for residents, have changed how residents spend their time, with only a minority of time spent in direct patient care when compared to time and motion studies a decade or more ago.⁴⁴

Conclusion

Duty hour limits are supported by the scientific literature, particularly limits on weekly hours and reducing the frequency of overnight call. However, many outcomes of interest are underrepresented in the empirical literature. Scientific rigor and established methods have required researchers to restrict their studies to a narrow subset of the large number of

variables of interest. These methods are challenged in properly addressing the multiple, often competing, attributes of the learning environment that collectively affect patient safety, resident learning, and resident well-being. Empirically derived limits must be examined for implementation and fit with context and stakeholder buy-in; all limits will benefit from a broader examination of their implementation that allows for a concurrent assessment of several outcomes of interest. Residents may play an important role in this. In addition to being affected, they are an energetic and insightful group, who may see the benefits and drawbacks of approaches in ways their program and institutional leaders and faculty may be blinded to.

References

- Philibert I, Friedmann P, Williams WT; ACGME Work Group on Resident Duty Hours. Accreditation Council for Graduate Medical Education. New requirements for resident duty hours. *JAMA*. 2002;288(9):1112–1114.
- 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(2):e3.
- Reed DA, Cook DA, Beckman TJ, et al. Association between funding and quality in published medical education research. *JAMA*. 2007;298(9):1002–1009.
- Friedman RC, Bigger JT, Kornfeld DS. The intern and sleep loss. *N Engl J Med*. 1971;285(4):201–203.
- Reuben DB. Depressive symptoms in medical house officers. Effects of level of training and work rotation. *Arch Intern Med*. 1985;145(2):286–288.
- Deaconson TF, O'Hair DP, Levy MF, et al. Sleep deprivation and resident performance. *JAMA*. 1988;260(12):1721–1727.
- Côté L, Turgeon J. Appraising qualitative research articles in medicine and medical education. *Med Teach*. 2005;27(1):71–75.
- Jacques CH, Lynch JC, Samkoff JS. The effects of sleep loss on cognitive performance of resident physicians. *J Fam Pract*. 1990;30(2):223–229.
- Hillson SD, Rich EC, Dowd B, et al. Call nights and patients care: effects on inpatients at one teaching hospital. *J Gen Intern Med*. 1992;7(4):405–410.
- Laine C, Goldman L, Soukup JR, et al. The impact of a regulation restricting medical house staff working hours on the quality of patient care. *JAMA*. 1993;269(3):374–378.
- Tanz RR, Charrow J. Black clouds. Work load, sleep, and resident reputation. *Am J Dis Child*. 1993;147(5):579–584.
- Yedidia MJ, Lipkin M Jr, Schwartz MD. Doctors as workers: work-hour regulations and interns' perceptions of responsibility, quality of care, and training. *J Gen Intern Med*. 1993;8(8):429–435.
- Petersen LA, Brennan TA, O'Neil AC, et al. Does housestaff discontinuity of care increase the risk for preventable adverse events? *Ann Intern Med*. 1994;121(11):866–872.
- Marcus CL, Loughlin GM. Effect of sleep deprivation on driving safety in housestaff. *Sleep*. 1996;19(10):763–766.
- Richardson GS, Wyatt JK, Sullivan JP, et al. Objective assessment of sleep and alertness in medical house staff and the impact of protected time for sleep. *Sleep*. 1996;19(9):718–726.
- Steele MT, MA OJ, Watson WA, et al. The occupational risk of motor vehicle collisions for emergency medicine residents. *Acad Emerg Med*. 1999;6(10):1050–1053.
- Parks DK, Yetman RJ, McNeese MC, et al. Day-night pattern in accidental exposures to blood-borne pathogens among medical students and residents. *Chronobiol Int*. 2000;17(1):61–70.
- Bellini LM, Baime M, Shea JA. Variation of mood and empathy during internship. *JAMA*. 2002;287(23):3143–3146.
- 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.
- 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.
- 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(2):125–134.
- Poulose BK, Ray WA, Arbogast PG, et al. Resident work hour limits and patient safety. *Ann Surg*. 2005;241(6):847–856; discussion 856–860.
- Gopal R, Glasheen JJ, Miyoshi TJ, et al. Burnout and internal medicine resident work-hour restrictions. *Arch Intern Med*. 2005;165(22):2595–2600.
- Rosen IM, Gimotty PA, Shea JA, et al. Evolution of sleep quantity, sleep deprivation, mood disturbances, empathy, and burnout among interns. *Acad Med*. 2006;81(1):82–85.
- Arora V, Dunphy C, Chang VY, et al. The effects of on-duty napping on intern sleep time and fatigue. *Ann Intern Med*. 2006;144(11):792–798.
- Hutter MM, Kellogg KC, Ferguson CM, et al. The impact of the 80-hour resident workweek on surgical residents and attending surgeons. *Ann Surg*. 2006;243(6):864–871; discussion 871–875.
- Singh H, Thomas EJ, Petersen LA, et al. Medical errors involving trainees: a study of closed malpractice claims from 5 insurers. *Arch Intern Med*. 2007;167(19):2030–2036.

28. Volpp KG, Rosen AK, Rosenbaum PR, et al. Mortality among hospitalized Medicare beneficiaries in the first 2 years following ACGME resident duty hour reform. *JAMA*. 2007;298(9):975–983.
29. Arora VM, Georgitis E, Siddique J, et al. Association of workload of on-call medical interns with on-call sleep duration, shift duration, and participation in educational activities. *JAMA*. 2008;300(10):1146–1153.
30. Jagannathan J, Vates GE, Pouratian N, et al. Impact of the Accreditation Council for Graduate Medical Education work-hour regulations on neurosurgical resident education and productivity. *J Neurosurg*. 2009;110(5):820–827.
31. Szymczak JE, Brooks JV, Volpp KG, et al. To leave or to lie? Are concerns about a shift-work mentality and eroding professionalism as a result of duty-hour rules justified? *Milbank Q*. 2010;88(3):350–381.
32. McCoy CP, Halvorsen AJ, Loftus CG, et al. Effect of 16-hour duty periods on patient care and resident education. *Mayo Clinic Proc*. 2011;86(3):192–196.
33. Desai SV, Feldman L, Brown L, et al. Effect of the 2011 vs 2003 duty hour regulation-compliant models on sleep duration, trainee education, and continuity of patient care among internal medicine house staff: a randomized trial. *JAMA Intern Med*. 2013;173(8):649–655.
34. Sen S, Kranzler HR, Didwania AK, et al. Effects of the 2011 duty hour reforms on interns and their patients: a prospective longitudinal cohort study. *JAMA Intern Med*. 2013;173(8):657–662; discussion 663.
35. Levine AC, Adusumilli J, Landrigan CP. Effects of reducing or eliminating resident work shifts over 16 hours: a systematic review. *Sleep*. 2010;33(8):1043–1053.
36. Henrich J, Heine SJ, Norenzayan A. The weirdest people in the world? *Behav Brain Sci*. 2010;33(2–3):61–83.
37. Wagner R, Weiss KB. Lessons learned and future directions: CLER National Report of Findings 2016. *J Grad Med Educ*. 2016;8(2 suppl 1):55–56.
38. Leape L, Berwick D, Clancy C, et al; Lucian Leape Institute at the National Patient Safety Foundation. Transforming healthcare: a safety imperative. *Qual Saf Health Care*. 2009;18(6):424–428.
39. Pennell NA, Liu JF, Mazzini MJ. Interns work hours. *N Engl J Med*. 2005;362(7):726–728.
40. Shea JA, Dinges DF, Small DS, et al. A randomized trial of a three-hour protected nap period in a medicine training program: sleep, alertness, and patient outcomes. *Acad Med*. 2014;89(3):452–459.
41. Nasca TJ. Open letter to the GME community. November 4, 2016. Chicago, IL: Accreditation Council for Graduate Medical Education. http://www.acgme.org/Portals/0/PDFs/Nasca-Community/NascaLetterToTheCommunity_11-4-16.pdf. Accessed November 11, 2016.
42. Philibert I, Nasca T, Brigham T, et al. Duty-hour limits and patient care and resident outcomes: can high-quality studies offer insight into complex relationships? *Annu Rev Med*. 2013;64:467–483.
43. 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.
44. Block L, Habicht R, Wu AW, et al. In the wake of the 2003 and 2011 duty hours regulations, how do internal medicine interns spend their time. *J Gen Intern Med*. 2013;28(8):1042–1047.



Ingrid Philibert, PhD, MBA, is Senior Vice President, Field Activities, Accreditation Council for Graduate Medical Education, and Executive Managing Editor, *Journal of Graduate Medical Education*.

Corresponding author: Ingrid Philibert, PhD, MBA, Accreditation Council for Graduate Medical Education, 401 N Michigan Avenue, Suite 2000, Chicago, IL 60611, 312.755.5003, iphilibert@acgme.org