Patient safety, resident well-being and continuity of care with different resident duty schedules in the intensive care unit: a randomized trial

Christopher S. Parshuram MB ChB DPhil, Andre C.K.B. Amaral MD, Niall D. Ferguson MD MSc, G. Ross Baker PhD, Edward E. Etchells MSc MD, Virginia Flintoft BN MSc, John Granton MD, Lorelei Lingard PhD, Haresh Kirpalani BM MSc, Sangeeta Mehta MD, Harvey Moldofsky MD, Damon C. Scales MD PhD, Thomas E. Stewart MD, Andrew R. Willan PhD, Jan O. Friedrich MD DPhil; for the Canadian Critical Care Trials Group

Author audio interview: soundcloud.com/cmajpodcasts/parshuram-resident. Author video summary: www.cmaj.ca/lookup/suppl/doi:10.1503 /cmaj.140752-/DC2

See related commentary, www.cmaj.ca/lookup/doi/10.1503/cmaj.150010

Abstract

Background: Shorter resident duty periods are increasingly mandated to improve patient safety and physician well-being. However, increases in continuity-related errors may counteract the purported benefits of reducing fatigue. We evaluated the effects of 3 resident schedules in the intensive care unit (ICU) on patient safety, resident well-being and continuity of care.

Methods: Residents in 2 university-affiliated ICUs were randomly assigned (in 2-month rotation-blocks from January to June 2009) to in-house overnight schedules of 24, 16 or 12 hours. The primary patient outcome was adverse events. The primary resident outcome was sleepiness, measured by the 7-point Stanford Sleepiness Scale. Secondary outcomes were patient deaths, preventable adverse events, and residents' physical symptoms and burnout. Continuity of care and perceptions of ICU staff were also assessed.

Results: We evaluated 47 (96%) of 49 residents, all 971 admissions, 5894 patient-days and 452 staff surveys. We found no effect of schedule (24-, 16- or 12-h shifts) on adverse events (81.3, 76.3 and 78.2 events per 1000 patient-days, respectively; $p = 0.7$) or on residents' sleepiness in the daytime (mean rating 2.33, 2.61 and 2.30, respectively; $p = 0.3$) or at night (mean rating 3.06, 2.73 and 2.42, respectively; $p = 0.2$). Seven of 8 preventable adverse events occurred with the 12-hour schedule ($p = 0.1$). Mortality rates were similar for the 3 schedules. Residents' somatic symptoms were more severe and more frequent with the 24-hour schedule ($p = 0.04$); however, burnout was similar across the groups. ICU staff rated residents' knowledge and decision-making worst with the 16-hour schedule.

Interpretation: Our findings do not support the purported advantages of shorter duty schedules. They also highlight the trade-offs between residents' symptoms and multiple secondary measures of patient safety. Further delineation of this emerging signal is required before widespread system change. **Trial registration:** ClinicalTrials.gov, no. NCT00679809.

Competing interests: None declared.

This article has been peer reviewed.

Correspondence to: Christopher Parshuram, christopher.parshuram@sickkids.ca

CMAJ 2015. DOI:10.1503 /cmaj.140752

Physician fatigue is common, is associated with worse physician well-being and more medical errors than for well-rested clinicians, and may compromise patient safety.[1–3] Shorter duty hours are purported to address these concerns,[4,5] but they necessitate more care transitions, which increases the risk of information loss.[6,7] The net effect on patient safety therefore depends on the relative balance between fatigue and continuity.[8–13] Currently, high-quality data to guide scheduling decisions are limited.

The complexity, acuity and therapeutic intensity of patients' conditions and their care make the intensive care unit (ICU) an ideal environment to evaluate the trade-offs between physician fatigue and continuity. Prior randomized studies have evaluated data for interns[14] or intensivists,[15] rather than the residents who provide most in-house overnight care in Canadian ICUs.[16] We evaluated the impact of 3 commonly used schedules[5] on patient safety and resident well-being.

Methods

We performed a randomized trial of resident schedules in 2 academic medical–surgical ICUs (at Mount Sinai Hospital and St. Michael's Hospital) in Toronto, Ontario. These ICUs admitted adult patients, accepted residents from accredited training programs and had 2-month ICU rotations with 5 or more residents per rotation. The residents were from internal medicine, anesthesia, surgery and emergency medicine training programs, and they performed in-house overnight duty with supervision from fellows and intensivists who were not required to be in-house. All patients admitted to participating ICUs during the study period were eligible for inclusion. The main outcomes were adverse events and resident fatigue. The study was conducted during six 2-month ICU rotations from January and June 2009.

Blocks of residents were randomly assigned to 1 of 3 schedules (see Appendix 1, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140752/-/DC1). Within each ICU, the unit of allocation was the 2-month rotation. Each schedule was applied for 1 rotation at each site, and each allowed for a maximum of 7 overnight shifts and provided 2 free weekends in 28 days. In one of the ICUs, the 24-hour schedule was intentionally allocated as the schedule for the first rotation. All other schedule allocations were random.

In the 24-hour schedule, the resident's duty period began at 8 am and finished at 8:30 am the next morning. This schedule is typical of Canadian ICUs[16] and is associated with disruption of circadian rhythm, acute sleep deprivation and physical stress.[1] With the 16-hour schedule, the shift began at 4:30 pm and finished at 8:30 am the next day. In both the 24-hour and 16-hour schedules, overnight duty was followed by 24 hours free of duty. Residents not working overnight worked from 8 am to 5 pm on weekdays; in the 16-hour schedule, one resident provided daytime coverage during weekends.

In the third schedule, 12-hour overnight duty, residents began at 8:30 pm and finished at 8:30 am the next day. Overnight duty was worked for 3 or 4 consecutive nights, followed by a 72-hour period free of clinical duties. Daytime duty involved short (8 am to 4:30 pm) or long (8 am to 8:30 pm) duty periods. The 12-hour schedule required one additional handover (8:30 to 9 pm) and, like the 16-hour schedule, required at least 5 residents. The schedules of other health care professionals working in each ICU were not modified.

In a representative week (from 8 am Monday to 8 am the following Monday), the mean number of hours worked per week was 59.4 hours in the 24-hour schedule, 53.2 hours in the 16-hour schedule and 52.4 hours in the 12-hour schedule.

Patient outcomes

The main patient outcome was adverse events, defined as any unplanned injury arising as a consequence of medical care during the ICU stay that was associated with morbidity, that required treatment, that prolonged the hospital stay or that resulted in disability at discharge.[17,18] Secondary patient outcomes were preventable adverse events, death in the ICU and severity of adverse events. Preventable adverse events were defined as adverse events that could have been avoided given current knowledge and standards of care.[17]

Adverse events were identified by daily prospective screening with a multimodal approach.[19] Trained nurses with 5–20 years of experience observed handover and multidisciplinary ward rounds, facilitated voluntary reporting and reviewed medical charts. Patients were followed until the earlier of 2 weeks after transfer from the ICU or hospital discharge, to capture adverse events first presenting after ICU discharge. The medical records of screen-positive patients were subjected to secondary physician review to determine whether an adverse event had occurred, to classify the event and to rate preventability. Physician reviewers (A.C.K.B.A. and J.O.F.) were intensivists with at least 5 years of ICU experience who were trained in review methodology and were unaware of schedule allocation at the time of their reviews.

Resident outcomes

The primary resident outcome — sleepiness — was assessed every 4 hours from 8 am until the end of scheduled duty on 4 randomly selected days per week using the Stanford Sleepiness Scale.[20,21] This scale is a widely used and highly validated measure of fatigue. Fatigue is an important aspect of well-being, with high face validity and empiric data linking it to other measures of well-being.[22]

Secondary resident outcomes were somatic symptoms and burnout. On 2 randomly selected days each week, residents rated the presence and severity of somatic symptoms experienced in the previous 24 hours. Symptoms of moderate or greater severity were identified. In the first and last weeks of each resident's 2-month rotation, psychologic well-being was assessed with the 3 domains of the Maslach Burnout Inventory.[23]

Continuity of care and ICU staff perceptions

The main measure of continuity of care was the number of prior days that residents and nurses had cared for an index patient. This variable was measured as part of the continuity survey[24] administered to the nurse and resident caring for 2 randomly selected patients each week. At the end of each 2-month ICU rotation, a survey was administered to ICU staff (i.e., health care professionals other than residents) to gauge residents' knowledge, clinical consequences of residents' activities, relationships of residents with other staff and impressions of resident fatigue.

Data collection

The Mount Sinai Hospital Research Ethics Board and the St. Michael's Hospital Research Ethics Board approved the protocol and waived the need for consent. We obtained consent from residents for individual-level data, but their consent was not required for participation in the allocated schedule. Completion of the continuity survey by health care professionals in the ICU implied consent.

The original study plan was a before-and-after study, with resident fatigue and continuity of care as primary outcomes. After starting the study but before analyzing the data, we made 2 changes. First, after allocating the first ICU to the 24-hour overnight schedule, we randomized the schedule allocations. Second, recognition that continuity of care would be challenging to operationalize prompted us to substitute adverse events as a co–primary outcome. The analytic plan was not changed.

We selected a 1-unit change on the Stanford Sleepiness Scale as a reasonable minimal difference for sample size calculations, based on data showing a 1.7- to 2.2-unit increase from before to after overnight duty.[25] With conservative assumptions, the study sample required six 2-month rotations with 6 residents per rotation. For adverse events, the 1800 admissions per schedule that we anticipated provided 80% power to achieve statistical significance with a relative rate difference of at least 0.26.

For each patient, we recorded age, urgency of admission and Acute Physiologic Assessment and Chronic Health Evaluation (APACHE) II score[26] in the first 24 hours in the ICU. On each day of admission, we recorded the utilization of 5 specific ICU therapies and the Multiple Organ Dysfunction Score[27] (Table 1).

Data analysis

We summarized demographic data using means and percentages. We used linear regression models to compare age at ICU admission, APACHE II score, maximum Multiple Organ Dysfunction Score and the arcsine-transformed proportion of ICU-days that each ICU therapy was used.[28] We reported adverse events and preventable adverse events per 1000 patient-days, with comparisons between schedules using unadjusted and APACHE-adjusted Poisson regression that accounted for clustering. The six 2-month clusters (ICU rotations) were used as random effects in the model. We presented the Stanford Sleepiness Scale measurements descriptively and used repeated-measures analysis of variance to evaluate data for day (8 am to 4 pm) and night (8 pm to 4 am). For each of the 3 domains of the Maslach Burnout Inventory, we identified the proportion of residents with high degrees of burnout and analyzed the absolute values[29] using linear regression with and without baseline adjustment. We analyzed data from continuity and ICU staff surveys using analysis of variance. We reported p values for 3-way comparisons across schedules and calculated 95% confidence intervals (CIs) for pairwise comparisons. We calculated response rates with conservative assumptions (Appendix 2, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140752/-/DC1). We conducted the analyses using SAS software, version 9.3 (SAS Institute Inc., Cary, NC).

Results

We included 807 patients with 971 ICU admissions, for a total of 5894 ICU-days (Figure 1; see also Appendix 3, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140752/-/DC1). The median age at ICU admission was 61 (interquartile range [IQR] 46–74) years, and the median APACHE II score was 19 (IQR 13–26). Patients spent a median of 3 (IQR 2–7) days in the ICU (Table 1). Forty-seven (96%) of 49 eligible residents consented to provide study measurements (Figure 1, Table 1). Most participating residents had 3 or fewer years of postgraduate experience (38/47 [81%]), were from internal medicine (23/47 [49%]) and were 30 years of age or younger (31/47 [66%]). Residents and nurses completed 217 continuity surveys (response rate 75.3% of 288 potential responses) and ICU staff completed 452 surveys (response rate 47.1% of 960 potential responses).

Patient outcomes

Screening identified 1602 potential adverse events. Physician reviewers verified 464 adverse events in 207 (25.6%) of the patients, equivalent to 78.7 adverse events per 1000 patient-days. The likelihood of health care causing harm was judged as "virtually certain" in 311 (67.0%) of the events; 93 events (20.0%)

Table 1: Patient and resident characteristics

Characteristic	All	24-h night	16-h night	12-h night	*p* value*
Patients					
No. admitted†	807	278	270	337	
Age, yr, median (IQR)	61 (46–74)	63 (50–76)	60 (45–73)	62 (45–73)	0.3
APACHE II score, median (IQR)	19 (13–26)	21 (14–27)	19 (13–25)	18 (12–25)	0.02
> 2 admissions in period, no. (%)	74 (9)	27 (10)	22 (8)	25 (7)	0.6
ICU days/patient, median (IQR)	3 (2–7)	3 (2–8)	4 (2–7)	3 (2–7)	0.2
No. of admissions	971	311	293	367	
Unplanned admissions, no. (%)	846 (87)	278 (89)	255 (87)	313 (85)	0.2
No. of patient-days in ICU	5894	2043	1651	2200	
Mechanical ventilation days, no. (per 1000 patient-days)	3607 (612.0)	1389 (679.9)	900 (545.1)	1318 (599.1)	0.02
HF oscillatory ventilation days, no. (per 1000 patient-days)	122 (20.7)	43 (21.0)	47 (28.5)	32 (14.5)	
Noninvasive ventilation days, no. (per 1000 patient-days)	406 (68.9)	171 (83.7)	151 (91.5)	84 (38.2)	0.001
Inotrope days, no. (per 1000 patient-days)	1322 (224.3)	471 (230.5)	408 (247.1)	443 (201.4)	0.7
Dialysis days, no. (per 1000 patient-days)	122 (20.7)	334 (163.5)	274 (166.0)	227 (103.2)	0.047
Maximum MODS, median (IQR)	6 (3–9)	6 (4–9)	5 (3–9)	5 (2–9)	0.4
Residents					
No. participating‡	47	15	15	17	
Training program,§ no. (%)					0.9
Internal medicine	23 (49)	6 (40)	7 (47)	10 (59)	
Anesthesia	9 (19)	3 (20)	3 (20)	3 (18)	
Surgery	8 (17)	3 (20)	3 (20)	2 (12)	
Age category, yr, no. (%)					0.1
25 to 30	31 (66)	10 (67)	9 (60)	12 (71)	
> 30 to 35	10 (21)	5 (33)	3 (20)	2 (12)	
> 35	4 (9)	0 (0)	2 (13)	2 (12)	
Postgraduate experience, yr, no. (%)					0.3
≤ 2	27 (57)	9 (60)	11 (73)	7 (41)	
3	11 (23)	4 (27)	1 (7)	6 (35)	
4 to 5	2 (4)	1 (7)	1 (7)	0 (0)	
> 5	5 (11)	1 (7)	1 (7)	3 (18)	
ICU experience, yr, no. (%) of residents					0.3
None	25 (53)	8 (53)	10 (67)	7 (41)	
< 1	18 (38)	6 (40)	3 (20)	9 (53)	
1 to < 2	2 (4)	1 (7)	1 (7)	0 (0)	
≥ 2	0 (0)	0 (0)	0 (0)	0 (0)	
No. of sleepiness measurements	1770	630	489	651	
Residents with sleepiness measurements, no. (%)	47 (100)	15 (100)	15 (100)	17 (100)	
Measurements per resident,¶ median (IQR)	34 (22–46)	39 (27–49)	36 (16–45)	30 (23–41)	0.4
Start-rotation Maslach Burnout Inventory,** no. (%)	45 (96)	15 (100)	14 (93)	16 (94)	
Depersonalization domain, mean score ± SD	11.4 ± 5.8	9.8 ± 4.9	11.4 ± 7.4	13.0 ± 4.8	0.3
No. (%) with high burnout by depersonalization	23 (51)	7 (47)	6 (40)	10 (59)	
Emotional exhaustion domain, mean score ± SD	25.4 ± 10.1	23.7 ± 10.2	26.4 ± 9.6	26.2 ± 11.0	0.7
No. (%) with high burnout by emotional exhaustion	18 (40)	5 (33)	7 (47)	6 (35)	
Personal accomplishment domain, mean score ± SD	36.5 ± 5.9	36.9 ± 7.4	35.3 ± 5.4	37.3 ± 4.9	0.6
No. (%) with high burnout by personal accomplishment	13 (29)	5 (33)	4 (27)	4 (24)	
Continuity survey, no. (% by row)					
All responses	217 (100)	74 (34)	61 (28)	82 (38)	
Primary nurse	123 (57)	41 (33)	34 (28)	48 (39)	
Primary resident	94 (43)	33 (35)	27 (29)	34 (36)	

Note: APACHE = Acute Physiology Assessment and Chronic Health Evaluation, HF = high-frequency, ICU = intensive care unit, IQR = interquartile range, MODS = Multiple Organ Dysfunction Score, SD = standard deviation.
*The reported *p* values are for each 3-way comparison, without adjustment.
†The total number of patients is less than the sum of patients in the 3 schedules, as some patients were admitted multiple times (possibly during more than one schedule type).
‡Two residents did not provide demographic information (one each from the 16- and 12-hour night schedules). The total number of residents in the group represented by each column was used as the denominator for column percentage calculations.
§Five residents were from other disciplines (3 in the 24-hour schedule, 1 in the 16-hour schedule and 1 in the 12-hour schedule).
¶Residents provided a minimum of 12 and a maximum of 83 sleepiness measurements.
**Higher scores on the emotional exhaustion and depersonalization domains and lower scores on the personal achievement domain indicate greater burnout. Two residents did not provide baseline measurements for the Maslach Burnout Inventory.

resulted in prolonged hospital stay, and 20 events (4.3%) resulted in disability lasting more than 1 month (Table 2). There were 166 adverse events (81.3 per 1000 patient-days) with the 24-hour schedule, 126 (76.3 per 1000 patient-days) with the 16-hour schedule and 172 (78.2 per 1000 patient-days) with the 12-hour schedule (Table 2). Pairwise comparisons for patient and resident outcomes with 95% CIs are presented in Appendix 4 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140752/-/DC1). There were no significant differences between schedules for rates of adverse events in unadjusted ($p = 0.7$) and adjusted ($p = 0.6$) analyses.

Preventable adverse events occurred 8 times in 7 (0.9%) of the patients, corresponding to 1.4 preventable adverse events per 1000 patient-days. Seven of the 8 preventable adverse events occurred with the 12-hour schedule. The regression model with the 16-hour schedule did not converge. Comparison of the 12- and 24-hour schedules was not significant before ($p = 0.1$) or after ($p = 0.4$) adjustment for APACHE II score. Four of the 8 preventable adverse events were associated with prolonged hospital stay, and all of these occurred with the 12-hour schedule. The overall ICU mortality rate was 16.5% (160/971), with rates of 18.3% (57/311) with the 24-hour schedule, 17.1% (50/293) with the 16-hour schedule and 14.4% (53/367) with the 12-hour schedule (unadjusted, $p = 0.2$; adjusted, $p = 0.6$).

Resident outcomes

The 47 residents provided a total of 1770 sleepiness measurements (including times when they were asleep), with a median of 34 (IQR 22–46) measurements per resident and a response rate of 56.7% (1770/3124 potential measurements) (Appendix 5, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140752/-/DC1). During the day, the mean rating on the Stanford Sleepiness Scale was 2.39 (standard deviation [SD] 1.13). This rating did not differ significantly between schedules ($p = 0.3$). At night, the mean rating was 2.72 (SD 1.37). Sleepiness increased from 8 pm to 4 am ($p < 0.001$) and was not statistically different between schedules ($p = 0.2$).

Residents reported their symptoms over the preceding 24 hours in 387 periods (response rate 74.6% of 519 potential reports). Symptoms were moderate or more severe in 113 (29.2%) of these reports (Appendix 6, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140752/-/DC1).

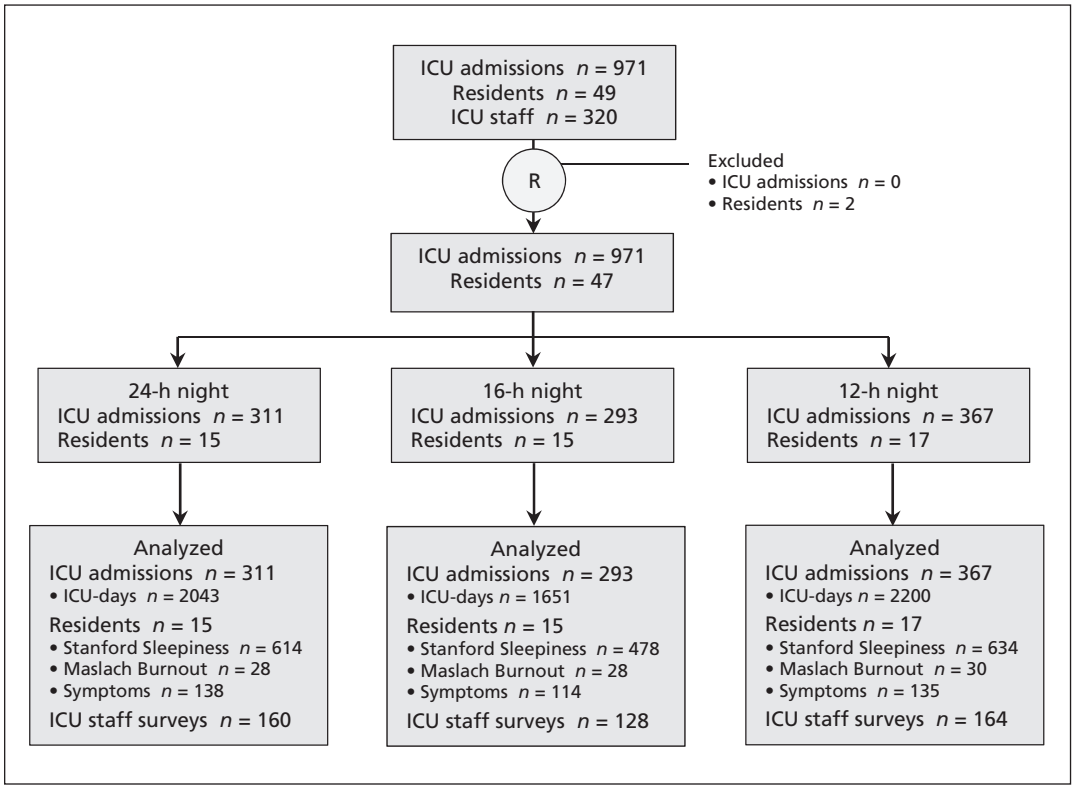

Figure 1: CONSORT diagram for a study of the effects of resident duty schedules in the intensive care unit (ICU). On the basis of the study design, we anticipated evaluating outcomes from 980 admissions for 5505 patient-days, 74 sleepiness and 13 symptom measurements per resident, 288 continuity results and 960 ICU staff survey responses. Two residents participated in the schedule but did not consent to providing data for analysis. The *n* values for the Maslach Burnout Inventory represent number of measurements.

Table 2: Patient and resident outcomes: overall and by overnight duty schedule

Outcome	Duty schedule; no. (per 1000 patient-days)*				*p* value
	All	24-h night	16-h night	12-h night	
Patients					
Adverse events†	464 (78.7)	166 (81.3)	126 (76.3)	172 (78.2)	0.7
Increased hospital stay	93 (15.8)	37 (18.1)	15 (9.1)	41 (18.6)	0.5
Disability					
None at discharge	151 (25.6)	59 (28.9)	39 (23.6)	53 (24.1)	0.9
Present < 1 mo after discharge	283 (48.0)	102 (49.9)	78 (47.2)	103 (46.8)	0.5
Present ≥ 1 mo after discharge	20 (3.4)	3 (1.5)	4 (2.4)	13 (5.9)	0.3
Errors‡					
Of commission	41 (7.0)	11 (5.4)	14 (8.5)	16 (7.3)	0.6
Of omission	246 (41.7)	85 (41.6)	67 (40.6)	94 (42.7)	> 0.9
Preventable adverse events†	8 (1.4)	1 (0.5)	0 (0)	7 (3.2)	0.1
ICU mortality, no. (%)	160 (16.5)	57 (18.3)	50 (17.1)	53 (14.4)	0.2
Residents					
Stanford Sleepiness Scale§					
Total no. of measurements (while awake)	1726	614	478	634	
Day (8 am, noon, 4 pm), no. of measurements	1286	468	350	468	
Median score (IQR)	2 (2–3)	2 (1–3)	2 (2–3)	2 (2–3)	
Mean score ± SD	2.39 ± 1.13	2.33 ± 1.20	2.61 ± 1.17	2.30 ± 0.99	0.3
Night (8 pm, midnight, 4 am), no. of measurements	440	146	128	166	
Median score (IQR)	2 (2–3)	3 (2–4)	2 (2–3)	2 (2–3)	
Mean score ± SD	2.72 ± 1.37	3.06 ± 1.57	2.73 ± 1.24	2.42 ± 1.21	0.2
At 4 am					
No. (%) of residents asleep¶	42 (28)	15 (32)	10 (21.3)	17 (29)	
No. of measurements on Stanford Sleepiness Scale	108	36	37	35	
Stanford Sleepiness Scale, mean score ± SD	3.74 ± 1.51	4.08 ± 1.66	3.81 ± 1.24	3.31 ± 1.55	0.2
*End-rotation Maslach Burnout Inventory***					
No. of residents providing data	41	13	14	14	
Depersonalization domain, mean score ± SD	11.9 ± 6.2	12 ± 6.9	12.2 ± 7.1	11.6 ± 5.0	> 0.9
No. (%) with high burnout by depersonalization	24 (59)	6 (46)	9 (64)	9 (64)	0.7
Emotional exhaustion domain, mean score ± SD	26.8 ± 10.3	28 ± 11.1	27.1 ± 9.5	25.4 ± 11.0	0.8
No. (%) with high burnout by emotional exhaustion	23 (56)	7 (54)	8 (57)	8 (57)	> 0.9
Personal accomplishment domain, mean score ± SD	34.6 ± 6.8	34.7 ± 6.6	33.1 ± 8.5	36.0 ± 5.4	0.6
No. (%) with high burnout by personal accomplishment	16 (39)	7 (54)	6 (43)	3 (21)	0.2
Somatic symptoms,†† no. of assessments	387	138	114	135	
Most severe symptoms, mean ± SD	2.08 ± 0.87	2.38 ± 0.91	1.86 ± 0.87	1.96 ± 0.77	0.05
Symptoms ≥ moderate severity, mean ± SD per report	0.62 ± 1.24	1.15 ± 1.71	0.38 ± 0.75	0.28 ± 0.73	0.04
Symptoms < moderate severity, no. (%) of assessments	274 (71)	80 (58)	84 (74)	110 (81)	
Continuity survey‡‡					
No. of responses from residents	90	32	26	32	
No. of days' care by residents, mean ± SD	2.34 ± 1.19	2.16 ± 1.19	2.73 ± 1.25	2.21 ± 1.10	0.1
No. of responses from nurses	123	41	34	48	
No. of days' care by nurses, mean ± SD	1.58 ± 0.87	1.78 ± 0.91	1.47 ± 0.86	1.48 ± 0.82	0.2

Note: ICU = intensive care unit, IQR = interquartile range, SD = standard deviation.
*Except where indicated otherwise.
†Adjustment for Acute Physiology Assessment and Chronic Health Evaluation (APACHE) score was performed for adverse events ($p = 0.6$), preventable adverse events ($p = 0.4$) and mortality ($p = 0.6$). We report results for the comparison of 12- and 24-hour schedules for evaluation of preventable adverse events, as models with the 16-hour schedule (zero events) did not converge. Adjustment for APACHE score did not change the p values for the occurrence and persistence of disability (all $p > 0.3$).
‡Adverse events were classified as errors of omission or errors of commission when the preventability rating was 2 or more on the 6-point scale.
§The Stanford Sleepiness Scale associates the following numeric values with descriptions of sleepiness: 1 = feeling active, vital, alert or wide awake; 2 = functioning at high levels, but not at peak, able to concentrate; 3 = awake but relaxed, responsive but not fully alert; 4 = somewhat foggy, let down; 5 = foggy, losing interest in remaining awake, slowed down; 6 = sleepy, woozy, fighting sleep, preferring to lie down; 7 = no longer fighting sleep, anticipating sleep onset soon, having dream-like thoughts. An additional category, "I was asleep at the time of assessment," was used for study coding purposes, but these ratings were excluded from analyses of the Stanford Sleepiness Scale.
¶Residents reported being asleep on 2 occasions other than at 4 am: once at midnight and once at noon.
**Overall (i.e., separate from rotations) post hoc comparison of start and end-rotation Maslach Burnout Inventory domain scores showed a significant decrease for the personal accomplishment domain (1.8, 95% confidence interval [CI] 3.5 to 0.03) and statistically insignificant increases for the emotional exhaustion domain (0.88, 95% CI –1.5 to 3.3) and the depersonalization domain (0.24, 95% CI –1.3 to 1.8).
††Residents who were on duty in the ICU indicated which of the following 14 symptoms they had experienced in the past 24 hours: headache, eye pain, blurred vision, light-headedness, anxiety, nausea, vomiting, diarrhea, epigastric pain, other abdominal pain, palpitations, chest pain, musculoskeletal pain and swollen feet. Symptom severity was rated as the worst of none = 1, mild = 2, moderate = 3, moderately severe = 4 and severe (constant distraction) = 5.
‡‡Each week, the nurse and resident providing care to each of 2 randomly selected patients completed the continuity survey. Data presented here are based on responses to the question, "In the last 7 days, how many days have you looked after this patient?"

Symptoms were absent or mild (i.e., no symptoms of moderate or greater severity) in 110 (81%) of the reports with the 12-hour schedule, 84 (74%) of those with the 16-hour schedule and 80 (58%) of those with the 24-hour schedule. Residents working the 24-hour schedule had greater severity of their worst symptom (p = 0.05) and reported more symptoms of at least moderate severity (p = 0.04) (Table 2).

The end-rotation Maslach Burnout Inventory scores indicated burnout in the emotional exhaustion domain for 23 (56%) of 41 residents, the depersonalization domain for 24 residents (59%) and the personal accomplishment domain for 16 residents (39%). There was no significant effect of schedule (all $p > 0.5$). The only factor significantly associated with end-rotation score was the initial score. Post hoc analysis showed that over the ICU rotation, the personal accomplishment domain decreased by a mean of 1.8 (95% CI 3.5–0.03) points (Table 2).

Continuity of care and ICU staff surveys

Residents completed 94 continuity questionnaires. Of the 90 responses to the question "In the last 7 days, how many days have you looked after this patient?" 23 (26%) indicated caring for the patient on that day only, 30 (33%) indicated looking after the patient on 1 other day, and 34 (38%) indicated looking after the patient for 3 or more days (Appendix 7, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140752/-/DC1). There were no differences between schedules (p = 0.1). In contrast, 73 (59%) of the 123 nurses who completed a continuity survey reported looking after the index patient for the first time on the day of the survey.

The 452 responses to the ICU staff survey included responses from 302 (66.8%) nurses, 58 (12.8%) respiratory therapists and 37 physicians (8.2%) (Appendix 8, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140752/-/DC1). ICU staff reported that residents working the 16-hour schedule were less familiar with the clinical (p = 0.002) and social (p = 0.03) details of their patients, and they perceived more suboptimal decisions by residents (p = 0.02). Residents working the 12-hour schedule were judged most alert overnight ($p < 0.001$). There were no differences in the support provided by front-line staff or supervising physicians or in perceptions of safety (Appendix 9, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140752/-/DC1).

Interpretation

In this randomized trial, we found no effect of resident schedules on adverse events and signals suggesting that safety might be compromised by shorter duty schedules. ICU staff perceived that residents knew fewer clinical and social details about patients and made lower-quality decisions when working a 16-hour schedule (Appendix 9), and 7 of the 8 preventable adverse events occurred with the 12-hour schedule.

Residents reported more somatic symptoms with the 24-hour schedule and sleepiness that was not significantly different across schedules. In 28% of sleepiness measurements taken at 4 am, the residents were asleep, and sleepiness varied among those who were awake overnight (Appendix 5). This variation may be related to interindividual variation in resistance to fatigue[30] and sleep or may be related to individuals' workload before 4 am.[1,31,32] With all 3 schedules, sleepiness was greatest at 4 am, consistent with the circadian rhythms of nocturnal sleepers and with the notion that time of day may be a more important determinant of fatigue than duration of duty.

Our findings that overnight duty periods of 12 or 16 hours may be somewhat better for residents and worse for patients are relevant in Canada, the United States and Europe, where these shorter schedules are increasingly used,[5] and underscore the need to further delineate this emerging signal before widespread system change.[33]

Our study can be added to 2 negative randomized studies that compared 32- and 16-hour duty periods for interns in the ICU[14] and in-house with home-call for critical care physicians.[15] In those studies, there were no differences in adverse events, preventable adverse events[14] or mortality rates.[14,15] Rates were similar to those observed in our study, except for the rate of preventable adverse events: 39 per 1000 patient-days[14] v. 1.4 per 1000 patient-days in our study. This difference in harmful errors may relate to system factors and the fact that we chose to assume best-practices were followed unless otherwise indicated. Two additional randomized controlled trials (RCTs) did not evaluate patient outcomes;[32,34] however, in one of these studies, the 14-hour overnight schedule was discontinued because of concerns about patient safety.[34]

We found that burnout, as assessed by the Maslach Burnout Inventory, was independent of schedule. Post hoc analyses showed a decrease in scores in the personal accomplishment domain over the ICU rotation, which suggests that effects on this burnout domain may be observable within a 2-month timeframe. Decreased burnout has been reported in before-and-after studies of duty-hour regulation, but those studies evaluated more residents over longer intervals.[35,36]

In previous RCTs, physician schedules were randomized within 3-week,[14] 7-day[15] and

4-week[32,34] periods. Two showed differences in resident sleepiness,[32,37] which suggests that 3–4 weeks is a sufficient duration to observe change in sleepiness (as indicated by electroencephalography or the Karolinska Sleepiness Scale). None of these randomized studies evaluated the Maslach Burnout Inventory.

Limitations

The limitations of this study included our inability to conceal the schedule, which enabled bias in subjective responses and screening for adverse events, as in previous studies.[14,32,34] However, the physician reviewers were unaware of schedule allocation during their reviews of adverse events. We endeavoured to account for unmeasured confounding by means of randomization, measurement of resident descriptors, evaluation of groups of moderate size and application of each schedule in both ICUs. We found differences in illness severity at the time of admission to the ICU and in therapies provided to the patient groups studied (Table 1), but statistical adjustment did not change our results.

Our study may have been underpowered. Significantly larger studies will be required to generate robust estimates of preventable adverse events and to describe smaller effects in other outcomes. Attribution of adverse events to individual residents might have increased the power. However, for the following reasons, we chose not to do this: it is the overall rate that is most relevant for patients, the origins of harmful errors are multifactorial,[38] and attribution would not reflect the complexities of ICU decision-making. The measurement of continuity was challenging, which led us to reprioritize it as a secondary outcome. Future work should address this limitation.

Finally, we did not formally monitor residents' adherence to their schedules, although informal review suggested adherence. Anecdotally, many residents disliked overnight duty in the 12-hour schedule, citing disconnection from the ICU team and loss of social contact with family and loved ones as a consequence of starting work at 8:30 pm for 3 or 4 consecutive nights. Future studies should evaluate these concerns.

Conclusion

We found no significant differences among 3 commonly used resident duty schedules in terms of adverse event rates and residents' sleepiness. No schedule protected against overnight fatigue or burnout. Our findings do not support the purported advantages of shorter duty and highlight trade-offs between residents' symptoms and multiple secondary measures of patient safety. More precise quantification will require the conduct of larger randomized studies.

References

1. Parshuram CS, Dhanani S, Kirsh JA, et al. Fellowship training, workload, fatigue and physical stress: a prospective observational study. *CMAJ* 2004;170:965-70.
2. Wu AW. Medical error: the second victim. The doctor who makes the mistake needs help too. *BMJ* 2000;320:726-7.
3. Veasey S, Rosen R, Barzansky B, et al. Sleep loss and fatigue in residency training: a reappraisal. *JAMA* 2002;288:1116-24.
4. Committee on Optimizing Graduate Medical Trainee (Resident) Hours and Work Schedules to Improve Patient Safety. *Resident duty hours: enhancing sleep, supervision and safety*. Washington: The National Academies Press; 2009.
5. National Steering Committee on Resident Duty Hours. *Fatigue, risk and excellence: towards a pan-Canadian consensus on resident duty hours*. Ottawa: Royal College of Physicians and Surgeons of Canada; 2013.
6. Borowitz SM, Waggoner-Fountain LA, Bass EJ, et al. Adequacy of information transferred at resident sign-out (in-hospital handover of care): a prospective survey. *Qual Saf Health Care* 2008;17:6-10.
7. Segall N, Bonifacio AS, Schroeder RA, et al. Can we make postoperative patient handovers safer? A systematic review of the literature. *Anesth Analg* 2012;115:102-15.
8. Parshuram CS. The impact of fatigue on patient safety. *Pediatr Clin North Am* 2006;53:1135-53.
9. Gandhi TK. Fumbled handoffs: one dropped ball after another. *Ann Intern Med* 2005;142:352-8.
10. Vidyarthi AR, Arora V, Schnipper JL, et al. Managing discontinuity in academic medical centers: strategies for a safe and effective resident sign-out. *J Hosp Med* 2006;1:257-66.
11. Fletcher KE, Parekh V, Halasyamani L, et al. Work hour rules and contributors to patient care mistakes: a focus group study with internal medicine residents. *J Hosp Med* 2008;3:228-37.
12. Asch DA, Parker RM. The Libby Zion case. One step forward or two steps backward? *N Engl J Med* 1988;318:771-5.
13. Jagsi R, Kitch BT, Weinstein DF, et al. Residents report on adverse events and their causes. *Arch Intern Med* 2005;165:2607-13.
14. 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;351:1838-48.
15. Kerlin MP, Small DS, Cooney E, et al. A randomized trial of nighttime physician staffing in an intensive care unit. *N Engl J Med* 2013;368:2201-9.
16. Parshuram CS, Kirpalani H, Mehta S, et al. In-house, overnight physician staffing: a cross-sectional survey of Canadian adult and pediatric intensive care units. *Crit Care Med* 2006;34:1674-8.
17. Baker GR, Norton PG, Flintoft V, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *CMAJ* 2004;170:1678-86.
18. Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *N Engl J Med* 1991;324:370-6.
19. Rothschild JM, Landrigan CP, Cronin JW, et al. The Critical Care Safety Study: the incidence and nature of adverse events and serious medical errors in intensive care. *Crit Care Med* 2005;33:1694-700.
20. Hoddes E, Zarcone V, Smythe H, et al. Quantification of sleepiness: a new approach. *Psychophysiology* 1973;10:431-6.
21. Stanford Sleepiness Scale. Palo Alto (CA): Stanford University School of Medicine; 1997. Available: http://web.stanford.edu/~dement/sss.html (accessed 2015 Jan. 28).
22. Pilcher JJ, Ott ES. The relationships between sleep and measures of health and well-being in college students: a repeated measures approach. *Behav Med* 1998;23:170-8.
23. Maslach C, Schaufeli WB, Leiter MP. Job burnout. *Annu Rev Psychol* 2001;52:397-422.
24. Pronovost P, Berenholtz S, Dorman T, et al. Improving communication in the ICU using daily goals. *J Crit Care* 2003;18:71-5.
25. Bartel P, Offermeier W, Smith F, et al. Attention and working memory in resident anaesthetists after night duty: group and individual effects. *Occup Environ Med* 2004;61:167-70.
26. Knaus WA, Draper EA, Wagner DP, et al. APACHE II: a severity of disease classification system. *Crit Care Med* 1985;13:818-29.
27. Marshall JC, Cook DJ, Christou NV, et al. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med* 1995;23:1638-52.
28. Snedecor GW, Cochran WG. *Statistical methods*. Ames (IA): Iowa State University Press; 1989.

29. Dyrbye LN, West CP, Shanafelt TD. Defining burnout as a dichotomous variable. *J Gen Intern Med* 2009;24:440, author reply 441.
30. Van Dongen HP, Bender AM, Dinges DF. Systematic individual differences in sleep homeostatic and circadian rhythm contributions to neurobehavioral impairment during sleep deprivation. *Accid Anal Prev* 2012;45(Suppl):11-6.
31. Bismilla Z, Breakey VR, Swales J, et al. Prospective evaluation of residents on call: before and after duty-hour reduction. *Pediatrics* 2011;127:1080-7.
32. Volpp KG, Shea JA, Small DS, et al. Effect of a protected sleep period on hours slept during extended overnight in-hospital duty hours among medical interns: a randomized trial. *JAMA* 2012;308:2208-17.
33. Ahmed N, Devitt KS, Keshet I, et al. A systematic review of the effects of resident duty hour restrictions in surgery: impact on resident wellness, training, and patient outcomes. *Ann Surg* 2014;259:1041-53.
34. 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;173:649-55.
35. Gopal R, Glasheen JJ, Miyoshi TJ, et al. Burnout and internal medicine resident work-hour restrictions. *Arch Intern Med* 2005; 165:2595-600.
36. Landrigan CP, Fahrenkopf AM, Lewin D, et al. Effects of the Accreditation Council for Graduate Medical Education duty hour limits on sleep, work hours, and safety. *Pediatrics* 2008;122:250-8.
37. Lockley SW, Barger LK, Ayas NT, et al. Effects of health care provider work hours and sleep deprivation on safety and performance. *Jt Comm J Qual Patient Saf* 2007;33(11 Suppl):7-18.
38. Reason J. Human error: models and management. *West J Med* 2000;172:393-6.

Affiliations: Department of Critical Care Medicine (Parshuram) and The Center for Safety Research at Child Health Evaluative Sciences (Parshuram, Willan), The Hospital for Sick Children, Toronto, Ont.; Interdepartmental Division of Critical Care Medicine (Parshuram, Amaral, Ferguson, Granton, Mehta, Scales, Friedrich,), Institute for Health Policy, Management and Evaluation (Baker, Flintoft), Department of Medicine (Etchells, Granton, Stewart), Department of Anaesthesia (Stewart) and Dalla Lana School of Public Health (Willan), University of Toronto, Toronto, Ont.; Department of Critical Care (Friedrich) and Li Ka Shing Knowledge Institute (Friedrich), St. Michael's Hospital, Toronto, Ont.; Sunnybrook Research Institute (Amaral, Scales), Department of Critical Care Medicine (Amaral, Scales) and Division of General Internal Medicine (Etchells), Sunnybrook Health Sciences Centre, Toronto, Ont.; Critical Care Medicine (Granton), University Health Network, Toronto, Ont.; Centre for Education Research and Innovation (Lingard) and Department of Medicine and Dentistry (Lingard), Western University, London, Ont.; Perelman School of Medicine (Kirpalani), University of Pennsylvania, Philadelphia, Pa.; Neonatology (Kirpalani), The Children's Hospital of Philadelphia, Philadelphia, Pa.; Department of Medicine, Division of Respirology and Critical Care Program (Ferguson, Mehta), Mount Sinai Hospital and University Health Network, Toronto, Ont.; Department of Medicine, (Mehta) Mount Sinai Hospital, Toronto, Ont.; Sleep Disorders Clinic of the Centre for Sleep and Chronobiology (Moldofsky), Toronto, Ont.; Department of Clinical Epidemiology and Biostatistics (Willan), McMaster University, Hamilton, Ont.; Research Institute (Ferguson), Toronto General Hospital, Toronto, Ont.

Contributors: Christopher Parshuram, Niall Ferguson, Ross Baker, Virginia Flintoft, John Granton, Lorelei Lingard, Haresh Kirpalani, Sangeeta Mehta, Harvey Moldofsky, Damon Scales, Thomas Stewart, Andrew Willan and Jan Friedrich, contributed to the conception of the study, and all authors contributed to the study design; Andre Amaral, Edward Etchells, Virginia Flintoff and Jan Friedrich contributed to acquisition of data; Christopher Parshuram and Andrew Willan were responsible for the analysis, and all authors contributed to the interpretation; Christopher Parshuram drafted the manuscript, and all authors contributed to revisions, approved the submitted version of the manuscript and agree to act as guarantors of the work.

Funding: The study was supported by funding from the Canadian Institutes of Health Research (reference number 84525), the Centre for Quality Improvement and Patient Safety of the University of Toronto and the Research Institute at the Hospital for Sick Children. These funding bodies had no role in the design or conduct of the study, including collection, management, analysis and interpretation of the data and preparation, review or approval of the manuscript. Jan Friedrich received funding to assist in completing a portion of the physician reviews of adverse events.

Data sharing: Patient data will be available on request for collaborative projects that are approved by the authors and begin at least 6 months after the date of publication.

Acknowledgments: The authors wish to recognize the following individuals for their contributions to the study: Orla Smith for local study facilitation; Amardeep Bali, Nadeene Blanchard, Kate Byrne, Nancy Cooper, Wendy Dolik, Susan Ferri, Helena Frndova, Jags Josan, Libby Kalman, Anne-Marie Krancevic, Kristen Middaugh, Kate Rosen, Simran Singh and Keeli Stith for management, collection, entry and review of data; and Nadeene Blanchard for preliminary analysis of data. All except Orla Smith were compensated for their efforts.