Medical Teacher

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/imte20

Impact of the resident duty hours on in-training examination score: A nationwide study in Japan

Kazuya Nagasaki, Yuji Nishizaki, Tomohiro Shinozaki, Hiroyuki Kobayashi, Taro Shimizu, Tomoya Okubo, Yu Yamamoto, Ryota Konishi & Yasuharu Tokuda

To cite this article: Kazuya Nagasaki, Yuji Nishizaki, Tomohiro Shinozaki, Hiroyuki Kobayashi, Taro Shimizu, Tomoya Okubo, Yu Yamamoto, Ryota Konishi & Yasuharu Tokuda (2021): Impact of the resident duty hours on in-training examination score: A nationwide study in Japan, Medical Teacher, DOI: 10.1080/0142159X.2021.2003764

To link to this article: https://doi.org/10.1080/0142159X.2021.2003764

MEDICAL TEACHER
https://doi.org/10.1080/0142159X.2021.2003764

Impact of the resident duty hours on in-training examination score: A nationwide study in Japan

Kazuya Nagasaki[a], Yuji Nishizaki[b], Tomohiro Shinozaki[c], Hiroyuki Kobayashi[a], Taro Shimizu[d], Tomoya Okubo[e], Yu Yamamoto[f], Ryota Konishi[g] and Yasuharu Tokuda[h]

[a]Department of Internal Medicine, Mito Kyodo General Hospital, University of Tsukuba, Ibaraki, Japan; [b]Division Medical Education, Juntendo University School of Medicine, Tokyo, Japan;[c]Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science, Tokyo, Japan; [d]Department of Diagnostic and Generalist Medicine, Dokkyo Medical University Hospital, Tochigi, Japan; [e]Research Division, National Center for University Entrance Examinations, Tokyo, Japan; [f]Division of General Medicine, Center for Community Medicine, Jichi Medical University School of Medicine, Tochigi, Japan; [g]Education Adviser Japan Organization of Occupational Health and Safety, Kanagawa, Japan; [h]Muribushi Okinawa for Teaching Hospitals, Okinawa, Japan

ABSTRACT

Purpose: The relationship between duty hours (DH) and the performance of postgraduate residents is needed to establish appropriate DH limits. This study explores their relationship using the General Medicine In-training Examination (GM-ITE).

Materials and methods: In this cross-sectional study, GM-ITE examinees of 2019 had participated. We analyzed data from the examination and questionnaire, including DH per week (eight categories). We examined the association between DH and GM-ITE score, using random-intercept linear models with and without adjustments.

Results: Five thousand five hundred and ninety-three participants (50.7% PGY-1, 31.6% female, 10.0% university hospitals) were included. Mean GM-ITE scores were lower among residents in Category 2 (45–50 h; mean score difference, −1.05; $p < 0.001$) and Category 4 (55–60 h; −0.63; $p = 0.008$) compared with residents in Category 5 (60–65 h; Reference). PGY-2 residents in Categories 2–4 had lower GM-ITE scores compared to those in Category 5. University residents in Category 1 and Category 5 showed a large mean difference (−3.43; $p = 0.01$).

Conclusions: DH <60–65 h per week was independently associated with lower resident performance, but more DH did not improve performance. DH of 60–65 h per week may be the optimal balance for a resident's education and well-being.

KEYWORDS
General medicine in-training examination; in-training examination; postgraduate medical education; resident duty hours; resident well-being

1. Introduction

In July 2018, the Japanese government enacted a bill related to 'Work Style Reform' to address the regular long working hours for all workers (MHLW 2018). In this reform, the annual maximum overtime was 720 h. In Japan, 8 h/day and 40 h/week are the legal working hour limits; working hours that exceed these limits are called overtime. These limits apply to all workers, including physicians and postgraduate residents. However, this new limit on overtime was exempted for doctors for 5 years in consideration of the profession's unique nature. A panel of experts from the Ministry of Health, Labour, and Welfare (MHLW) compiled recommendations for how to regulate overtime for physicians beginning in 2024 and continuing until 2035. They set a typical physician's overtime hours at 960 h per year (approximately duty hour [DH] of 60 h per week) (MHLW 2019, 2020). However, overtime for residents who require intensive education was tentatively set at 1,860 h per year (approximately total DH of 80 h per week). The statement

Practice points

- The maximum duty hours (DH) for postgraduate residents in Japan will be limited to about 80 h per week from 2024, but there is an insufficient assessment of the possible impact on the education and health of residents.
- Previous literature suggested that duty hour restrictions (DHR) on residents in other countries may improve residents' well-being but may also negatively impact their education.
- This study shows that residents in Japan need at least 60–65 h of DH per week in their clinical training to acquire a certain level of clinical knowledge. However, clinical knowledge does not proportionately increase with increases in DH.
- Policies should support a more balanced view of duty hour restrictions as less duty hours is not always better.

CONTACT Yuji Nishizaki ynishiza@juntendo.ac.jp Division of Medical Education, Juntendo University School of Medicine, 2-1-1 Hongo Bunkyo-ku, Tokyo 113-8421, Japan

Supplemental data for this article can be accessed here.

mentioned that longer overtime hours for residents could increase the health hazards of residents. At the same time, there were concerns that future reductions in resident DH could harm physician training. However, there is scant evidence on the relationship between the educational outcome of residency training and DH in Japan, and the optimal duty hour is unclear.

Japanese residents work long hours, and this has been linked to burnout, depression, excessive stress, and even suicide. A 2018 MHLW survey found that physicians in their 20 s (mainly residents) worked the longest hours, an average of 76.1 h per week (MHLW 2017). Recent surveys found that 18%–33.3% of residents in Japan suffered from burnout (Miyoshi et al. 2016; Nishimura et al. 2019). About one-half of PGY-1 residents who worked more than 100 h per week showed severe depressive symptoms (Ogawa et al. 2018). Excessive DH of residents is a global problem, as it has been reported that residents in the United States (US) and the United Kingdom commonly work more than 60 h per week (Baldwin et al. 1997; Barger et al. 2005). Currently, the DH limit for residents is 80 h and 48 h per week in the US and European Union, respectively (IOM 2009).

Long DH and resident educational outcomes have been seldom investigated in Japan as well as worldwide. The impact of an 80-h work hour restriction (WHR) on resident education and health outcomes has been extensively reported in the US. Extensive surveys of residents have reported that a WHR improved resident burnout and well-being but negatively affected patient care and education (Goitein et al. 2005; Gopal et al. 2005). Meanwhile, one survey suggested that a WHR reduced the number of patient care experiences but did not lessen important clinical experiences (Jagsi et al. 2006). In a systematic review of WHRs for surgical residents, the effects on resident health were inconsistent across studies, but lower-quality education and certification exam scores, and patient safety concerns were identified (Ahmed et al. 2014). Overall, WHR could improve the well-being of residents but negatively affect their education (Mansukhani et al. 2012). These studies revealed that WHR can worsen educational outcomes. However, they are insufficient in determining the appropriate DH level.

With this trend, it is important to examine the relationship between resident DH and educational outcomes to determine optimal DH for residents. Therefore, we used the General Medicine In-training Examination (GM-ITE), an in-training examination of residents in Japan, as well as a training environment questionnaire to assess the relationship between DH and clinical knowledge acquisition of residents.

2. Materials and methods

2.1. Study populations

We conducted a multicenter cross-sectional observational study of postgraduate year 1 (PGY-1) and postgraduate year 2 (PGY-2) residents. After completing a 6-year medical school education in Japan, medical students enter a 2-year postgraduate residency program. Residents must undergo supervised training and rotate through seven specialties (internal medicine, emergency medicine, community-based medicine, surgery, anesthesiology, pediatrics, psychiatry, and obstetrics and gynecology). The MHLW regulates training programs and sets goals for acquiring basic knowledge and general medicine skills for residents. A computerized national matching system is in place, allowing students to apply to hospitals of their choice with clinical training facilities across the country (Kozu 2006; Teo 2007). After a 2-year residency program, most residents enter specialty-based residency training.

Our participants were all residents who underwent GM-ITE at the end of the 2019 academic year (from January 21 to January 28 of 2020). After the examination, participants were immediately asked about average DH during their residency program as part of a questionnaire assessing their residency training and work environment. All participants read and signed the informed consent document before the survey. The Ethics Review Board of Juntendo University School of Medicine approved the study.

2.2. Measurements

In our study, the primary exposure variable was average DH per week, which was calculated as the sum of weekday work duty, night emergency department (ED) duty, and weekend work duty (Supplementary Material). Standby time during ED was also included in DH. The duration of DH was self-reported in the following eight categories: Category 1 (<45 h), Category 2 (45–50 h), Category 3 (50–55 h), Category 4 (55–60 h), Category 5 (60–65 h), Category 6 (65–70 h), Category 7 (70–80 h), and Category 8 (>80 h). The Category of DH was defined as the lower limit of no or minimal overtime and the upper limit of 80 h/week, which is the new resident DH limit set by the MHLW.

The GM-ITE was an in-training examination of clinical knowledge using a methodology similar to that of the US Residency Internal Medicine In-Training Examination (IM-ITE) (Garibaldi et al. 2002; Kanna et al. 2009; Perez and Greer 2009). The GM-ITE aimed to provide medical residents and program directors with an objective, reliable, valid assessment of clinical knowledge. The GM-ITE contained 60 questions categorized into four major topics: clinical diagnosis, physical examination/procedure, medical interview/professionalism, and subspecialties. A part of the GM-ITE questions incorporates video and audio questions. The GM-ITE was first introduced in 2011 by the Japan Organization of Advancing Medical Education (JAMEP, a non-profit organization), and it is administered continuously every year. The examination was developed annually by a committee of experienced physicians and peer-reviewed by an independent committee. Residents who belonged to training hospitals that have applied for the GM-ITE were eligible to take the examination. To the best of our knowledge, the results of this exam are not currently used in the recruitment process for specialty training.

2.3. Statistical analyses

The hypothesized positive association between resident DHs and GM-ITE scores was examined using a random-intercept linear model. Hospital variability was considered as normal random intercepts. Considering that 60 h per week is the basic upper limit of DH for all doctors, we set Category 5 (60–65 h) as a reference for the analyses. We excluded those who failed to provide information about

Table 1. Characteristics of residents categorized by duty hours.

	C1: <45 h N = 134	C2: 45–50 h N = 514	C3: 50–55 h N = 827	C4: 55–60 h N = 832	C5: 60–65 h N = 1048	C6: 65–70 h N = 754	C7: 70–80 h N = 677	C8: >80 h N = 807
Demographics								
Men (%)	64.9	63.8	68.9	67.9	70.4	66.7	66.6	72.0
PGY-2 (%)	45.5	48.1	51.3	52.0	50.7	53.1	51.3	49.1
Hospital types (%)								
University	14.9	15.2	13.4	8.1	9.5	9.7	9.3	6.2
University branch	7.5	7.4	5.0	5.6	4.4	5.4	5.0	2.6
Community	77.6	77.4	81.6	86.3	86.1	84.9	85.7	91.2
ED duties per month (%)								
None	18.7	7.4	3.5	2.0	2.7	2.0	2.7	1.5
1–2	23.9	27.4	21.2	12.4	10.2	9.4	8.9	6.1
3–5	46.3	60.7	70.3	78.2	76.5	75.5	72.1	64.8
6 or more	10.4	3.7	4.6	6.5	9.0	12.5	15.7	26.5
Unknown	0.0	0.8	0.5	0.7	1.4	0.7	0.6	0.7
Assigned inpatients (%)								
0–4	40.3	29.8	23.5	19.0	17.6	14.7	13.0	10.5
5–9	47.8	58.6	62.3	65.5	61.9	63.0	60.0	54.4
0–14	10.4	8.2	9.8	10.5	14.6	15.0	17.9	18.2
15 or more	0.7	1.9	1.7	1.6	2.5	3.6	6.1	13.9
Unknown	0.7	1.6	2.7	3.1	3.1	3.7	3.1	2.9
Self-study time per day (%)								
None	7.5	3.3	3.4	4.2	3.4	3.8	4.3	6.8
0–30 min	54.5	49.2	39.1	37.7	35.7	32.5	31.6	33.6
31–60 min	24.6	38.1	42.6	42.4	41.5	47.7	43.4	38.0
61–90 min	10.4	8.4	12.6	12.9	16.2	13.3	16.5	17.1
91 min or more	2.2	1.0	2.4	2.8	3.1	2.7	4.0	4.5

C1–C8: Category 1 to Category 8; h: hours; PGY: postgraduate year; ED: emergency department.
Note. A positive trend between duty hours and each category: men (%), community hospital (%), number of ED duties, inpatient caseload, and self-study time. Calculation of the average duty hours was based on the sum of weekday work duty, night emergency department (ED) duties, and weekend work duty.

DH. We repeated the analyses by PGY and hospital types (community hospitals, university hospitals, and university-affiliated hospitals). Our regression models were also adjusted for sex, PGY, hospital types, the number of ED duties, the number of assigned inpatients, and self-study time. GM-ITE scores among residents in Japan are associated with the number of ED duties, the number of assigned inpatients, and self-study time (Shimizu et al. 2013; Kinoshita et al. 2015; Mizuno et al. 2016; Nishizaki et al. 2017). The number of ED duties, assigned inpatients, and self-study time was self-reported in the following categories:

- Number of ED duties: 0, 1–2, 3–5, 6 or more, and unknown.
- Number of assigned inpatients: 0–4, 5–9, 10–14, 15 or more, and unknown.
- Self-study time per day: none, 1–30, 31–60, 61–90, and 91 min or more.

All analyses were conducted using SAS version 9.4 (Cary, NC, USA) and all linear mixed-effect models with hospital-random intercepts were fitted via a restricted maximum likelihood estimation (the default setting in the MIXED procedure) assuming independent errors within hospitals. This study followed STROBE guidelines.

3. Results

A total of 6,869 residents from 539 teaching hospitals participated in the examination, with a voluntary questionnaire response rate of 89.7% (6,164/6,869). We excluded 571 residents who did not provide DHs from the analysis. A total of 5,593 participants were included in the study (2,839 PGY-1 and 2,754 PGY-2). A summary of the baseline characteristics of the participants categorized by DHs is shown in Table 1. Of the participants, 1,770 (31.6%) were female. Among the training hospitals, 562 (10.0%) were university hospitals, 278 (5.0%) were university-affiliated hospitals, and 4,753 (85.0%) were community hospitals. The mean (±*SD*) GM-ITE score of all participants was 29.4 ± 2.0. As a result of trend analyses, the number of ED duties ($p < 0.001$), inpatient caseload ($p < 0.001$), and self-study time ($p < 0.001$) exhibited associations with DHs. A summary of the participants' baseline characteristics categorized by PGY and hospital types is shown in Supplementary Material.

The results of the random-intercept linear model analysis for residents are presented in Figures 1–3 and Supplementary Material. Multicollinearity of each variable is also assessed in Supplementary Material. Mean GM-ITE scores were lower among residents in Category 1 (mean score difference (MSD), −1.35; 95% CI, −2.28 to −0.41; $p = 0.005$), Category 2 (−1.36; −1.91 to −0.80; $p < 0.001$), and Category 4 (−0.69; −1.16 to −0.22; $p < 0.004$) compared with residents in Category 5 (Reference). When analyzed by PGY, PGY-1 residents in Category 1 (−1.79; −3.12 to −0.47; $p = 0.01$) and Category 2 (−0.79; −1.54 to −0.03; $p = 0.04$) had lower GM-ITE scores than those in Category 5. Meanwhile, GM-ITE scores were lower among PGY-2 residents in Categories 1–4 than in Category 5 (Figure 2(A)). University hospital residents in Category 1 and Category 5 showed a large mean difference (−4.12; −6.67 to −1.58; $p < 0.002$) (Figure 3(A)).

After adjusting for sex, PGY, hospital types, monthly ED duties, number of assigned inpatients, and self-study time, mean GM-ITE scores were lower among residents in Category 2 (MSD, −1.05; CI, −1.64 to −0.52; $p < 0.001$) and Category 4 (−0.63; −1.10 to −0.16; $p = 0.008$) compared with residents in Category 5 (Figure 1(B)). PGY-1 residents in Category 1 (−1.45; −2.80 to −0.11; $p = 0.03$) had lower GM-ITE scores compared to those in Category 5.

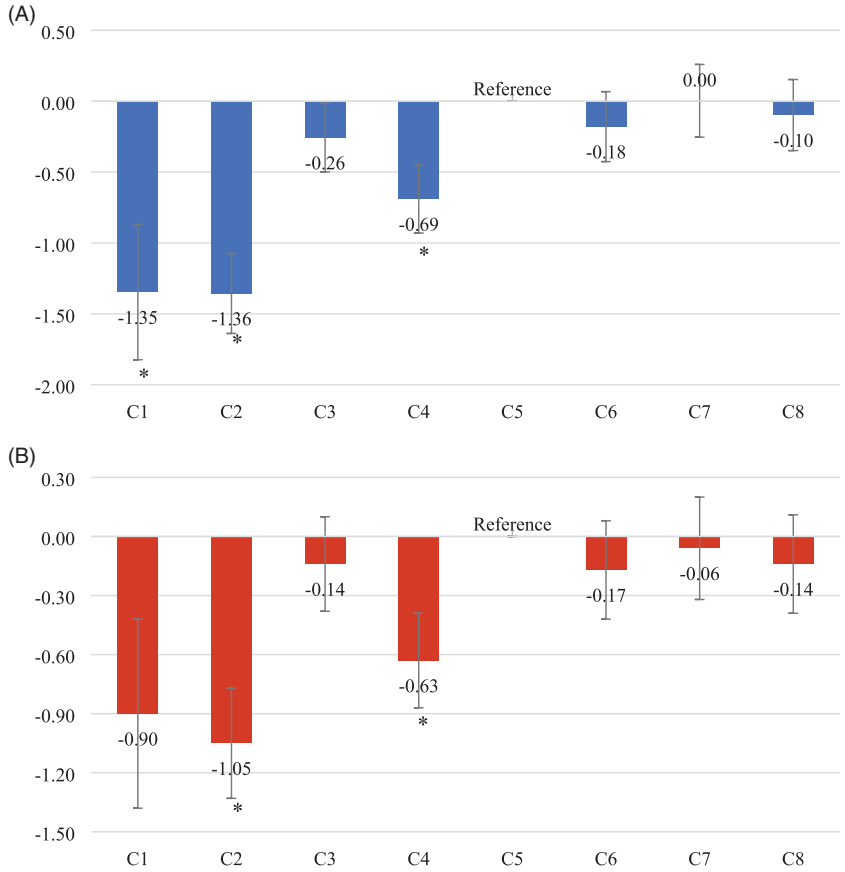

Figure 1. Estimates of mean score differences on GM-ITE between residents of each category and Category 5. (A) Univariable Analysis. (B) Multivariable Analysis†. *Note*. The residents with a DH of <60 h per week had lower GM-ITE scores than those with a DH of 60–65 h per week. However, those with a DH of >65 h per week did not have higher scores. Resident duty hours per week comprised eight categories: category 1 (<45), category 2 (45–50), category 3 (50–55), category 4 (55–60), category 5 (60–65), category 6 (65–70), category 7 (70–80), and category 8 (>80). †Adjusted for gender, hospital types, postgraduate years, monthly ED duties, number of assigned inpatients, and self-study time. Non-responders of the number of ED duties and inpatients were included in 'unknown.' Non-responders of self-study time were excluded from the multivariable analysis. * Indicates $p < 0.05$. Error bars indicate standard errors estimated in the random-intercept model. C1–C8: Category 1 to Category 8; h = hours; GM-ITE: general medicine in-training examination.

Furthermore, GM-ITE scores were lower among PGY-2 residents in Categories 2–4 than in Category 5 (Figure 2(B)); University hospital residents in Category 1 and Category 5 continued to show a large mean difference (−3.43; −6.04 to −0.83; $p = 0.01$) (Figure 3(B)).

The random-intercept multivariable-adjusted model estimated the variance of random effects (i.e. hospital-level variation of the score) as 2.55 (standard error: 0.35) and the variance of residuals (i.e. resident-level variation within the hospital) as 25.15 (standard error: 0.50). This indicates the presence of (1) a hospital-level variation in the GM-ITE score and (2) a small to moderate similarity within each hospital, which may be quantified by an intraclass correlation coefficient of $2.55/(2.55 + 25.15) = 0.092$.

4. Discussion

Overall, postgraduate residents with DH of <60 h per week had lower GM-ITE scores than those with DH of 60–65 h per week. In comparison, residents with DH of more than 65 h per week did not have higher GM-ITE scores. In residents with DH of <60 h per week, there was a greater decrease in scores for PGY-2 than PGY-1. In addition, in those groups, residents in university hospitals had a greater decrease in scores than those in other hospital types. These differences remained after multivariable adjustment. This study showed an independent association between resident DH and GM-ITE score.

In groups with less DHs, PGY-2 had particularly lower GM-ITE scores than PGY-1. In a previous report, PGY-2 had generally higher GM-ITE scores than PGY-1, possibly reflecting the number of experienced cases (Mizuno et al. 2016). One hypothesis is that the DHs of PGY-2 residents reflect the number of experienced cases but not those of PGY-1 residents. The reason may be that the maximum number of cases assigned to PGY-1 is less than that to PGY-2 or that most of the duty hours in PGY-1 are not direct patient care. McCoy et al. reported that the IM-ITE performance of PGY-3 residents reflects the number of case encounters but not that of PGY-2 residents (McCoy et al. 2013). Our results may be comparable to their study. Another point regarding PGY differences is that residents with more DHs appeared to have higher scores for PGY-1 than for PGY-2. Hence, working more than 60–65 h/week in PGY-2 may be detrimental from an educational perspective compared with that in PGY-1. Since these results were not statistically significant in this study, future research is needed.

There are two possible reasons for lower scores in the residents with less DHs in university hospitals compared to other hospital settings. First, residents who choose university hospitals may be less interested than residents in community hospitals in experiencing a large number of clinical cases (Teo 2007). Another possibility is that residents in university hospitals are research- and specialty-oriented

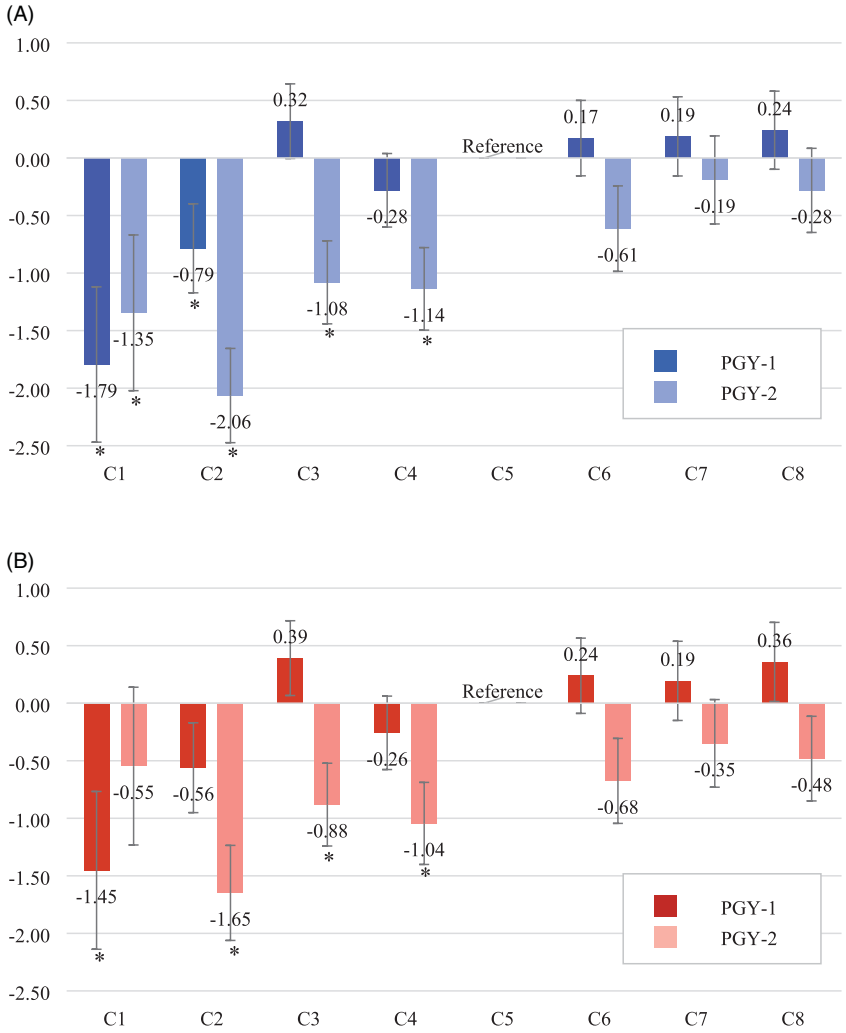

Figure 2. Estimates of mean score differences on GM-ITE between residents of each category and Category 5 stratified by postgraduate year. (A) Univariable analysis. (B) Multivariable analysis†. *Note.* In residents with a DH of <60 h per week, there was a greater decrease in the scores of PGY-2 residents than in those of PGY-1 residents. Resident duty hours per week comprised eight categories: category 1 (<45), category 2 (45–50), category 3 (50–55), category 4 (55–60), category 5 (60–65), category 6 (65–70), category 7 (70–80), and category 8 (>80). †Adjusted for gender, hospital types, monthly ED duties, number of assigned inpatients, and self-study time. Non-responders of the number of ED duties and inpatients were included in 'unknown.' Non-responders of self-study time were excluded from the multivariable analysis. * Indicates $p < 0.05$. Error bars indicate standard errors estimated in the random-intercept model. C1–C8: Category 1 to Category 8; h: hours; GM-ITE: general medicine in-training examination.

and may be less motivated to learn about primary care, as measured by the GM-ITE score. One of the main goals of clinical training in Japan is acquiring comprehensive clinical competence. Even clinical residents at university hospitals need to have adequate case experience and the associated clinical knowledge (Yano et al. 1992).

This study shows that residents in Japan need at least 60–65 h of DH per week (maximum overtime of 960–1200 h per year) in their clinical training to acquire a certain level of clinical knowledge. However, it is important to note that clinical knowledge does not proportionately increase with increases in DH. Therefore, to maintain residents' well-being, there should be an appropriate upper limit on residents' DH. As mentioned earlier, beginning in 2024, the maximum overtime for residents is1860 h per year from an educational perspective (MHLW 2019, 2020). However, our study results indicated that 1860 h/year of overtime for residents seems not optimal and needs to be reduced. In addition, although this study evaluated the relationship between duty hours and GM-ITE scores, it did not evaluate the amount of opportunity to learn in their DHs. Increasing the educational time during training is also important to improve educational outcomes as well as to maintain resident health. Each training program should encourage residents to self-study by providing online educational resources and constructing an in-hospital learning environment (Edson et al. 2010).

This study has several limitations. First, DH is self-reported in the survey. With self-reporting, residents may over- or under-report the number of hours worked, and the trend may vary by Category group. Although one study indicates that resident self-reporting is relatively accurate, it is preferable to use an automated attendance management system to measure hours (Chadaga et al. 2012). Two, this study does not evaluate residents' mental health, such as burnout and depression. Even if scores are higher in groups that work more than 60–65 h, health outcomes may worsen. These outcomes need to be assessed together to determine the optimal resident DH. Three, the GM-ITE is a written test. Based on its nature, it can only assess knowledge and associated cognitive skills. Therefore, the GM-ITE score does not necessarily guarantee real-world clinical performance. The current study does not adequately assess physical examination and clinical skills, communication, and professionalism, which are difficult to evaluate only via a written test. Fourth, this study assesses DHs in terms of

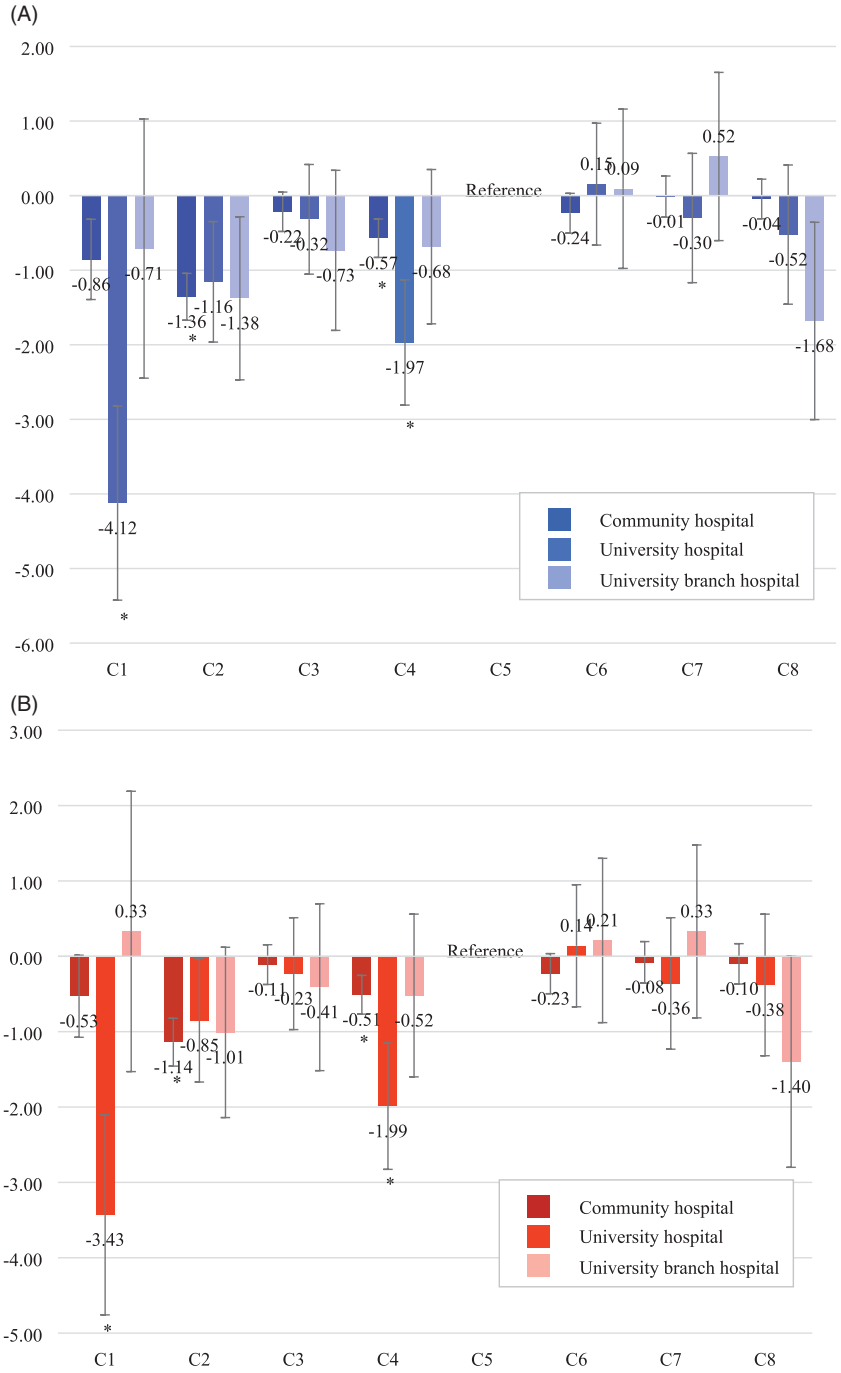

Figure 3. Estimates of mean score differences on GM-ITE between residents of each category and Category 5 stratified by hospital type. (A) Univariable analysis. (B) Multivariable analysis†. *Note*. In residents with a DH of <60 h per week, residents in university hospitals had a greater score decrease than those in other hospital types. Resident duty hours per week comprised eight categories: category 1 (<45), category 2 (45–50), category 3 (50–55), category 4 (55–60), category 5 (60–65), category 6 (65–70), category 7 (70–80), and category 8 (>80). †Adjusted for gender, postgraduate years, monthly ED duties, number of assigned inpatients, and self-study time. Non-responders of the number of ED duties and inpatients were included in 'unknown.' Non-responders of self-study time were excluded from the multivariable analysis. * Indicates $p < 0.05$. Error bars indicate standard errors estimated in the random-intercept model. C1–C8: Category 1 to Category 8; h: hours; GM-ITE: general medicine in-training examination.

the average number of hours worked over 2 years, ignoring each rotation's differences. In several training hospitals, night emergency duty continues regardless of the department, and work during weekends and holidays are consistent. However, the length of weekday and weekend duty hours varies per department. The MHLW study found that working hours in each department in Japan differed by a maximum of 15 h per week (MHLW 2017). It is necessary to know the DHs for each rotation to assess residents' DH and performance more accurately. Fifth, there is a potential for selection bias in this study. There are about 18,000 residents in Japan, and only about one-third participated in the study. The proportions of participants in terms of sex in this study are similar to all Japanese residents. However, the proportion of residents affiliated with university hospitals is lower. Notably, the percentages of residents in universities in Japan are 40.7% in 2018 and 38.9% in 2019. Meanwhile, the percentage of participants in this study was 10.0%. Although about 23 (30%) of 81 university hospitals participated in this study, it will weaken the generalization of the study results. In addition, because participation in the GM-ITE is voluntary for training hospitals, participating hospitals may have a higher quality of education than hospitals that do not participate. Sixth, this study does not evaluate a resident's baseline clinical knowledge acquisition. Baseline GM-ITE scores need to be adjusted to

improve the measurement of the impact of DH on resident training.

5. Conclusions

In conclusion, we show that GM-ITE scores decrease among postgraduate residents with DH of <60 h per week. Scores do not increase among those with DH of more than 65 h per week compared to those with DH of 60–65 h per week. This study provides essential data for setting an appropriate upper limit of DH for residents in Japan.

Ethical approval

The Ethics Review Board of Juntendo University School of Medicine approved the study.

Acknowledgments

We thank members of JAMEP for their assistance. The authors would like to thank Enago (www.enago.jp) for the English language review.

Disclosure statement

Dr. Nishizaki received an honorarium from the JAMEP as the GM-ITE project manager. Dr. Tokuda and Dr. Okubo are the JAMEP director. Dr. Nagasaki received an honorarium from the JAMEP as a reviewer of GM-ITE. Dr. Kobayashi received an honorarium from the JAMEP as a speaker of the JAMEP lecture. Dr. Shimizu and Dr. Yamamoto received an honorarium from the JAMEP as exam preparers of GM-ITE. Dr. Nishizaki, Dr. Tokuda, Dr. Okubo, Dr. Nagasaki, Dr. Kobayashi, Dr. Shimizu, and Dr. Yamamoto were not involved in the analysis.

Glossary

Duty hours: All clinical and academic activities related to the program, including patient care (both inpatient and outpatient), administrative duties relative to patient care, the provision for transfer of patient care, time spent in-house during call activities, and scheduled activities, such as conferences. Duty hours do not include reading and preparation time spent away from the duty site.[1]

Note

1. ACGME website. [accessed 2021 Jul 21]. https://www.acgme-i.org/Accreditation-Process/Glossary/.

Funding

The author(s) reported there is no funding associated with the work featured in this article.

Author contributions

Dr. Shinozaki had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Nishizaki and Tokuda: study concept and design. Shinozaki, Shimizu, Yamamoto, and Konishi: acquisition, analysis, or interpretation of data. Nagasaki: manuscript drafting. Nishizaki and Kobayashi: critical revision of the manuscript for important intellectual content. Shinozaki: statistical analysis. Nishizaki and Tokuda: administrative, technical, or material support. Nishizaki, Kobayashi, and Tokuda: supervision.

Prior presentations

The results of the study were reported on 5 November 2020, at the 'Committee on Promoting Reform of the Working Style of Physicians' organized by the Ministry of Health, Labour and Welfare [https://www.mhlw.go.jp/stf/newpage_14763.html (in Japanese)].

Notes on contributors

Kazuya Nagasaki, MD, works as a staff physician at Mito Kyodo General Hospital.

Yuji Nishizaki, MD, MPH, PhD, is an associate professor of medical education at Juntendo University School of Medicine.

Tomohiro Shinozaki, MPH, PhD, is a Jr. associate professor of biostatistics at Tokyo University of Science.

Hiroyuki Kobayashi, MD, PhD, is a professor of medicine at the University of Tsukuba and works as a head of the internal medicine department in Mito Kyodo General Hospital.

Taro Shimizu, MD, PhD, MPH, MBA, is a professor of medicine at Dokkyo Medical University Hospital.

Tomoya Okubo, PhD, is an associate professor in the department of test analysis and evaluation at the National Center for University Entrance Examinations.

Yu Yamamoto, MD, is an assistant professor of the division of general medicine at Jichi Medical University.

Ryota Konishi, MD, MS, works as an education advisor at the Japan Organization of Occupational Health and Safety.

Yasuharu Tokuda, MD, MPH, is a director of the Muribushi Okinawa Center for Teaching Hospitals.

ORCID

Kazuya Nagasaki http://orcid.org/0000-0002-8312-7802
Yuji Nishizaki http://orcid.org/0000-0002-6964-6702
Tomohiro Shinozaki http://orcid.org/0000-0003-3395-9691
Hiroyuki Kobayashi http://orcid.org/0000-0001-8377-0091
Taro Shimizu http://orcid.org/0000-0002-3788-487X
Yu Yamamoto http://orcid.org/0000-0003-2247-4963
Ryota Konishi http://orcid.org/0000-0001-5538-1507
Yasuharu Tokuda http://orcid.org/0000-0002-9325-7934

Data availability statement

Due to the nature of this research, participants of this study did not agree for their data to be shared publicly, so supporting data is not available.

References

Ahmed N, Devitt KS, Keshet I, Ahmed N, Devitt KS, Keshet I, Spicer J, Imrie K, Feldman L, Cools-Lartigue J, et al. 2014. A systematic review of the effects of resident duty hour restrictions in surgery: impact on resident wellness, training, and patient outcomes. Ann Surg. 259(6):1041–1053.

Baldwin PJ, Dodd M, Wrate RW. 1997. Young doctors' health-I. How do working conditions affect attitudes, health and performance? Soc Sci Med. 45(1):35–40.

Barger LK, Cade BE, Ayas NT, Cronin JW, Rosner B, Speizer FE, Czeisler CA, Harvard Work Hours, Health, and Safety Group. 2005. Extended work shifts and the risk of motor vehicle crashes among interns. N Engl J Med. 352(2):125–134.

Chadaga SR, Keniston A, Casey D, Albert RK. 2012. Correlation between self-reported resident duty hours and time-stamped parking data. J Grad Med Educ. 4(2):254–256.

Edson RS, Beckman TJ, West CP, Aronowitz PB, Badgett RG, Feldstein DA, Henderson MC, Kolars JC, McDonald FS. 2010. A multi-

institutional survey of internal medicine residents' learning habits. Med Teach. 32(9):773–775.
Garibaldi RA, Subhiyah R, Moore ME, Waxman H. 2002. The in-training examination in internal medicine: an analysis of resident performance over time. Ann Intern Med. 137(6):505–510.
Goitein L, Shanafelt TD, Wipf JE, Slatore CG, Back AL. 2005. The effects of work-hour limitations on resident well-being, patient care, and education in an internal medicine residency program. Arch Intern Med. 165(22):2601–2606.
Gopal R, Glasheen JJ, Miyoshi TJ, Prochazka AV. 2005. Burnout and internal medicine resident work-hour restrictions. Arch Intern Med. 165(22):2595–2600.
[IOM] Institute of Medicine (US). 2009. Committee on optimizing graduate medical trainee (resident) hours and work schedule to improve patient safety. In: Ulmer C, Miller Wolman D, Johns MME, editors. Resident duty hours: enhancing sleep, supervision, and safety. Washington: National Academies Press.
Jagsi R, Shapiro J, Weissman JS, Dorer DJ, Weinstein DF. 2006. The educational impact of ACGME limits on resident and fellow duty hours: a pre–post survey study. Acad Med. 81(12):1059–1068.
Kanna B, Gu Y, Akhuetie J, Dimitrov V. 2009. Predicting performance using background characteristics of international medical graduates in an inner-city university-affiliated internal medicine residency training program. BMC Med Educ. 9(1):42.
Kinoshita K, Tsugawa Y, Shimizu T, Tanoue Y, Konishi R, Nishizaki Y, Shiojiri T, Tokuda Y. 2015. Impact of inpatient caseload, emergency department duties, and online learning resource on general medicine in-training examination scores in Japan. Int J Gen Med. 8: 355–360.
Kozu T. 2006. Medical education in Japan. Acad Med. 81(12): 1069–1075.
Mansukhani MP, Kolla BP, Surani S, Varon J, Ramar K. 2012. Sleep deprivation in resident physicians, work hour limitations, and related outcomes: a systematic review of the literature. Postgrad Med. 124(4):241–249.
McCoy CP, Stenerson MB, Halvorsen AJ, Homme JH, McDonald FS. 2013. Association of volume of patient encounters with residents' in-training examination performance. J Gen Intern Med. 28(8): 1035–1041.
[MHLW] Ministry of Health, Labour and Welfare. 2018. Act on the arrangement of related acts to promote work style reform. [accessed 2020 Sep 24]. https://www.mhlw.go.jp/content/000332869.pdf.
[MHLW] Ministry of Health, Labour and Welfare. 2020. Addressing the reform of the working style of physicians in the clinical training system. [accessed 2020 Sep 24]. https://www.mhlw.go.jp/content/10803000/000590866.pdf
[MHLW] Ministry of Health, Labour and Welfare. 2019. Report of the committee for the reform of the working style of physicians. [accessed 2020 Sep 24]. https://www.mhlw.go.jp/content/10800000/000496522.pdf
[MHLW] Ministry of Health, Labour and Welfare. 2017. Survey on physicians' working conditions and working style intentions. [accessed 2020 Sep 24]. https://www.mhlw.go.jp/file/05-Shingikai-10801000-Iseikyoku-Soumuka/0000161146.pdf
Miyoshi R, Matsuo H, Takeda R, Komatsu H, Abe H, Ishida Y. 2016. Burnout in Japanese residents and its associations with temperament and character. Asian J Psychiatr. 24:5–9.
Mizuno A, Tsugawa Y, Shimizu T, Nishizaki Y, Okubo T, Tanoue Y, Konishi R, Shiojiri T, Tokuda Y. 2016. The impact of the hospital volume on the performance of residents on the general medicine in-training examination: a multicenter study in Japan. Intern Med. 55(12):1553–1558.
Nishimura Y, Miyoshi T, Obika M, Ogawa H, Kataoka H, Otsuka F. 2019. Factors related to burnout in resident physicians in Japan. Int J Med Educ. 10:129–135.
Nishizaki Y, Mizuno A, Shinozaki T, Okubo T, Tsugawa Y, Shimizu T, Konishi R, Yamamoto Y, Yanagisawa N, Shiojiri T, et al. 2017. Educational environment and the improvement in the general medicine in-training examination score. J Gen Fam Med. 18(5): 312–314.
Ogawa R, Seo E, Maeno T, Ito M, Sanuki M, Maeno T. 2018. The relationship between long working hours and depression among first-year residents in Japan. BMC Med Educ. 18(1):50.
Perez JA Jr., Greer S. 2009. Correlation of United States medical licensing examination and internal medicine in-training examination performance. Adv Health Sci Educ. 14(5):753–758.
Shimizu T, Tsugawa Y, Tanoue Y, Konishi R, Nishizaki Y, Kishimoto M, Shiojiri T, Tokuda Y. 2013. The hospital educational environment and performance of residents in the general medicine in-training examination: a multicenter study in Japan. Int J Gen Med. 6: 637–640.
Teo A. 2007. The current state of medical education in Japan: a system under reform. Med Educ. 200741(3):302–308.
Yano E, Yamaoka K, Sugita S, Kobayashi Y, Niino N, Fukui T, Yamakado M, Nishizaki O, Ogata T, Segami K. 1992. Comparing postgraduate medical education at university and non-university hospitals in Japan. Acad Med. 67(1):54–58.