Self-Reported Subjective Workload of On-Call Interns

Kathlyn E. Fletcher, MD, MA
Alexis M. Visotcky, MS
Jason M. Slagle, PhD
Sergey Tarima, PhD
Jeff Whittle, MD, MPH
Matthew B. Weinger, MD
Marilyn M. Schapira, MD, MPH

Abstract

Background Workload has traditionally been measured by using surrogates, such as number of patients admitted or census, but these may not fully represent the complex concept of workload.

Objective We measured self-reported subjective workload of interns and explored the relationship between subjective workload and possible predictors of it.

Methods Trained research assistants observed internal medicine interns on call on a general medicine service. Approximately once an hour, the research assistants recorded the self-reported subjective workload of the interns by using Borg's Self-Perceived Exertion Scale, a 6 to 20 scale, and also recorded their own perceptions of the intern's workload. Research assistants continuously recorded the tasks performed by the interns. Interns were surveyed before and after the observation to obtain demographic and census data.

Results Our sample included 25 interns, with a mean age of 28.6 years (SD, 2.4 years). Mean self-reported subjective workload was 12.0 (SD, 2.4). Mean self-reported subjective workload was significantly correlated with intern age ($r = 0.49$, $P < .05$), but not with team or intern census, number of admissions, or number of patients cross-covered. Self-reported subjective workload in the period after sign-out was significantly higher than in the period before and during sign-out ($P < .001$).

Conclusions Self-reported subjective workload was not associated with traditional measures of workload. However, receiving sign-out and assuming the care of cross-coverage patients may be related to higher subjective workload in interns. Given the patient safety implications of workload, it is important that the medical education community have tools to evaluate workload and identify contributors to it.

Kathlyn E. Fletcher, MD, MA, is Associate Professor, Milwaukee VA Medical Center (VAMC) and Department of Internal Medicine, Medical College of Wisconsin; **Alexis M. Visotcky, MS,** is Biostatistician, Department of Biostatistics, Medical College of Wisconsin; **Jason M. Slagle, PhD,** is Assistant Professor, Department of Anesthesia, Vanderbilt University; **Sergey Tarima, PhD,** is Biostatistician, Department of Biostatistics, Medical College of Wisconsin; **Jeff Whittle, MD, MPH,** is Professor, Milwaukee VAMC and Department of Internal Medicine, Medical College of Wisconsin; **Matthew B. Weinger, MD,** is Professor, Tennessee Valley VAMC and Department of Anesthesia, Vanderbilt University; and **Marilyn M. Schapira, MD, MPH,** is Associate Professor, Philadelphia VAMC and University of Pennsylvania Perelman School of Medicine.

Funding: This study was funded by the VA Health Services Research & Development grant No. PPO 09-259-01. Dr Weinger's effort on this project was also supported by Geriatric Research Education and Clinical Center at the VA Tennessee Valley Healthcare System.

This study was presented as a poster at the National Society of General Internal Medicine Meeting, Orlando, FL, May 2012.

The authors wish to thank Amanda Grippen, Nathan Meyer, Andrew Kordus, Kimberly Spencer, Jessica Schmidt, Sherrie Smaxwill, Eric Porterfield, and Lance Trautman for their contributions to this project.

Corresponding author: Kathlyn E. Fletcher, MD, MA, 5000 W National Avenue, PC Division, Milwaukee, WI 53295, 414.384.2000 ex. 46450, kathlyn.fletcher@va.gov

Received August 17, 2012; revision received February 12, 2013; accepted March 18, 2013.

DOI: http://dx.doi.org/10.4300/JGME-D-12-00241.1

Introduction

Driven to a considerable extent by the duty hour limitations of 2003[1] and 2011,[2] resident schedules have changed dramatically. Traditional concerns about fatigue[3,4] and burnout[5] have been at least partially replaced with new concerns about experience[6] and workload.[7]

Workload and work intensity are related concepts, and both have been studied to a limited degree in physicians. Workload has been defined in the human factors literature as being the extent to which a person has used the cognitive and physical resources needed for task performance.[8] Workload is usually conceptualized in medicine as the number of patients being cared for or the amount of work that needs to be done. Work intensity is more focused on the perceived experience of the worker with respect to the difficulty of a specific task or encounter.[9] Work intensity can be thought of as a moderator of workload, so a fixed amount of work will require more cognitive resources (workload) if the intensity is higher.

Physician workload seems to be important. In a 1993 survey study, residents reported that workload contributed to many serious patient care mistakes.[10] A focus group

study corroborated that workload was still an important factor in such mistakes after the 2003 implementation of duty hour limits.[11] Work intensity has an impact on physicians as well. In a study of Swiss physicians, work intensity was found to be correlated with emotional exhaustion and aversion to patients.[12]

Previous work in anesthesia used a scale from ergonomics[13] to measure the self-reported subjective workload (ie, work intensity) of physicians and nurses in the operating room.[14] Building on these methods, we designed the following study. This was part of a larger study of the work of interns while on call. The aim of this part of the study was to examine the self-reported subjective workload of internal medicine interns as they worked during an on-call shift.

What is known

Traditional measures of physician workload, such as patients admitted or census, may not fully reflect the complexity of workload.

What is new

A study assessed whether self-reported subjective workload was associated with traditional measures of workload.

Limitations

Study was conducted prior to 2011 duty hour standards implementation; small sample size, and single-site study limit generalizability.

Bottom line

Self-reported workload was not associated with traditional measures of physician workload, yet may assist in better allocation of resident work, contributing to enhanced patient safety.

Methods

Study Design

This was a prospective observational study with a cross-sectional survey component. Some of the methods have been previously described.[15]

Setting

This study was conducted at the Milwaukee VA Medical Center (VAMC) between May and October 2010. During the study, there were 4 general internal medicine ward teams. Each team had 1 second- or third-year resident, 2 interns, and 3 medical students (1 in the fourth year and 2 in the third year). Teams admitted patients from approximately 3 PM to 7 AM in an every-fourth-day cycle. Most of the work during this time was without direct attending supervision. One intern went home at 9 PM and returned the next morning. The other intern and resident stayed in the hospital overnight. This overnight intern also cross-covered the patients of the other general medicine teams after they had left for the day.

Subjects

All interns who rotated on the general medicine ward services during the study period were eligible for inclusion. There were no exclusion criteria. We followed the participating interns on a call day when they were scheduled to be the overnight intern. Each participating intern was shadowed once. Written informed consent was obtained, and the interns were given a $20 gift card.

Measures

We used the Borg Rating of Perceived Exertion Scale[13] to capture the self-reported subjective workload of interns. This instrument was originally designed to capture subjective workload while exercising, and its scale (6 to 20) reflected the workload that corresponded to heart rates between 60 and 200.[13,16,17] The scale has been used to measure self-reported subjective workload in physicians[18–20] and has been validated as an overall measure of workload in anesthesiologists, while they are attending to patients in the operating room (which includes tasks that are cognitive as well as physical).[18,21] We adapted it by adding our own domain-relevant verbal anchors (FIGURE) based on previous work to determine the difficulty of various tasks performed by interns while on call.[22] Specifically, at 6, we used the verbal anchor of "Calm" with the description of "Sleeping or surfing the Internet." At 20, we used the verbal anchor of "Intense" with the example of "Running a code." The midpoint was "Obtaining a history."

Physician Demographics and Brief Survey Items

We surveyed the interns to ascertain age, sex, and number of months in training. We also asked about other

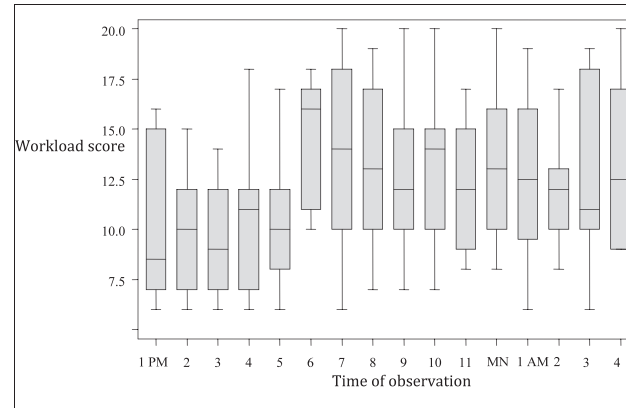

FIGURE **Self-reported Subjective Workload Ratings by Hour**

The time on the x-axis indicates the beginning of that observation hour, so that the 1 PM bar includes ratings between 1 PM and 1:59 PM. The workload scores are based on the Borg Rating of Perceived Exertion Scale (6–20 scale, with 6 = least exertion or "calm"). The outermost bars represent the 95% confidence intervals, the shaded rectangles represent the interquartile ranges, and the bars inside the rectangles represent the means.

parameters of workload. On the morning of the observation, the interns were asked (1) team census at the start of the observation day and (2) individual intern census at the start of the observation day. After the last team signed out, interns were asked to record the number of patients they covered for other teams overnight (''cross-cover load''). Finally, on the postcall morning, after admitting was over, interns were asked the number of patients that they admitted during the call period. One person recorded a value of 10 admissions, which is highly unlikely in our system, and this individual's data were not included in analyses about total number of admissions.

Data Collection Procedures

The purpose of data collection was 2-fold: (1) to record the work done by the interns and (2) to elicit the self-reported subjective workload of the interns as they performed their work. Interns' self-reported subjective workload is the focus of this report. Trained research assistants (RAs) shadowed the interns as they moved through the hospital. Data collection began at approximately 1 PM and ended at 5 AM the following morning. Times were chosen to maximize the chance of capturing the work that was done while the interns were admitting new patients and cross-covering old patients. The RAs did observe direct patient encounters, but left for sensitive parts (such as the physical examination). The RAs did not directly observe the interns while they were asleep.

The RAs used laptop computers with customized software for collecting data. This task analysis and workload measurement software is effective at measuring the effects of a variety of performance-shaping factors.[18,21,23,24] We programmed the software to display 59 possible tasks performed by an intern while on call (eg, examining a patient, writing orders), which were lumped into the following work categories: direct patient care, clinical computer work, nonpatient communication, teaching and learning, transit, and downtime.[22] The RA continuously recorded each of the intern's tasks by clicking on the appropriate task button. Each of the tasks was rigorously defined. The software automatically logged the time each task was initiated. The software allowed the RAs to capture multitasking. The RAs asked the interns for clarification on tasks, as needed.

Of particular interest to this study, the software also sent data collection prompts to the RAs at random time intervals between 30 and 90 minutes to assess perceived workload, using the Borg scale described above. The RA rated his or her perception of the intern's workload first (ie, observer-rated workload) and then had the intern rate his or her own workload (ie, self-reported workload). The RA held up a sign for the intern that said, ''What is your workload right now?'' The scale with anchors was printed under the question. If an intern was asleep, they were not woken up to answer the workload question.

We trained the RAs to be accurate and reliable observers. During their training, they observed teams rounding and physicians on call, and practiced identifying the tasks of a physician while working in the hospital. We trained the RAs not to interact with the teams, other than briefly assessing their workload, during the observations so that they would not influence the work of the interns. The interns did not know which tasks were being recorded.

The Milwaukee VAMC Institutional Review Board and Research and Development Committees approved this project.

Data Analysis

The data from each observation was saved in a text file with each line containing a task, the task start time, and time-stamped annotations. Data were transferred into Microsoft Excel (Microsoft Corp, Redmond, WA) files so that task times flowed continuously. When RA observers overlapped, we used the data from the observer taking over the shift. We then audited and cleaned the data. Our data analysis program was written in Visual Basic for Applications (Microsoft Corp). We used this program to analyze each intern's data file by calculating and collating the minutes spent on each task.

We combined the tasks into the 6 work categories and calculated descriptive statistics across cases, including means and standard deviations of the time spent on each task and work category. Means and standard deviations were calculated for the interns' self-rated workload across the shift and for the observers' ratings of the interns' workload. Statistical analysis was done by using SAS 9.2 (SAS Institute Inc, Cary, NC).

We used Pearson correlations to examine the univariate relationships between mean workload across the observation period, other parameters of workload (cross-cover load, intern census at start of day), and intern demographic factors, and to examine the relationship between mean intern-rated and observer-rated workload. We graphically examined the mean workload over time by looking at the workload in 1-hour increments (ie, 1–2 PM, 2–3 PM, etc). We used mixed linear modeling to compare the mean workload in 2 time periods: (1) before and during the hours when other teams were signing out to the intern subjects; and (2) after the last sign-out was given to the intern. Monte-Carlo Markov Chain with 20 000 repetitions was used to estimate *P* values. We also used analysis of variance to compare the mean workload associated with the 3 most

common types of tasks performed: direct patient care, clinical documentation work, and communication with other health care workers. All tests were 2-tailed, with significance at the $P < .05$ level.

Results

We had a participation rate of 69% (25 of 36), with a mean age of 28.6 (SD, 2.4) years and 56% (14) being women.[15] Participants reported a mean of 4 (SD, 3.7) months in training. Interns started the day with a mean of 2.6 (SD, 1.6) patients, with a range of 0 to 5 patients. The mean team census at the start of the day was 7.2 (SD, 2.7), with a range of 2 to 12. They reported a mean of 3.9 (SD, 1.8) admissions. The mean number of patients cross-covered overnight was 27.7 (SD, 12.1), with a range of 6 to 50 patients.

Self-reported subjective workload was measured a mean of 9.2 (SD, 2.0) times per intern, while observer-rated workload was measured a mean of 9.3 (SD, 1.9) times. Mean self-reported subjective workload was 12.0 (SD, 2.4), and mean observer-rated workload was 11.6 (SD, 1.8). These were highly correlated, ($r = 0.92$, $P < .01$). Mean self-reported subjective workload was significantly correlated with intern age ($r = 0.49$, $P < .05$), but not with the number of months in training ($r = -0.14$, $P = .50$). None of the traditional parameters of workload were significantly associated with self-reported subjective workload: intern's census at the start of the day ($r = 0.21$, $P = .31$), team census at the start of the day ($r = 0.14$, $P = .51$), number of patients cross-covered ($r = 0.18$, $P = .44$), or number of patients admitted ($r = 0.01$, $P = .32$).

Self-reported subjective workload varied over the course of the observation period, with an increase noted around 6 PM (FIGURE). Because sign-out often occurred around this time, we explored the effect of sign-out on workload by fitting the data with a linear mixed model. In our model, we included 2 fixed effects: (1) the fixed effect of assessor (intern or observer) and (2) the fixed effect of the time period (before/during versus after sign-out); and 2 random effects: (1) the intern effect and (2) the effect of the paired (intern self-rated and observer-rated) assessment of the workload. The effect of assessor was not significant ($P = .10$). The workload in the period after sign-out was significantly higher than the workload in the period before and during which sign-outs took place ($P < .001$), with a model-based estimated increase equal to 1.89 units on the workload scale. Interns' self-reported subjective workload did not vary depending on whether they were performing tasks related to direct patient care, computer work, or communication with other health care workers at the time of the rating (TABLE).

TABLE **Mean Workload by Task Performed Immediately Before Workload Assessment**

Workload Score by Category of Work	No. of Observations	Mean (SD)
Direct patient care	12	12.5 (3.0)
Clinical computer work	114	12.1 (3.5)
Communication with other health care workers	74	12.3 (3.9)

Discussion

The workload of physicians is a potential contributor to patient safety, but traditional methods for measuring workload are limited to the number of patients admitted or cared for by a given physician. In this study, we reported on the self-reported subjective workload of internal medicine interns during an on-call period. Mean workload over the observation period was 12.0, with variation over the course of the shift. Workload after the nonovernight teams had signed out was significantly higher than before. We were unable to demonstrate a relationship between the type of task being performed and the self-reported subjective rating of workload. We also found no relationship between the traditional parameters of workload and mean self-reported subjective workload.

It is interesting and intuitive that self-reported subjective workload increased after sign-out of the other teams' patients had occurred since that corresponds with interns becoming responsible for more patients. However, the absolute number of patients cross-covered did not correlate with overall workload. These 2 findings may mean that there is a threshold at which additional patients no longer add to the subjective workload of interns. We also considered alternative hypotheses to explain the increase in subjective workload during that time period. One explanation is that the rate of patient admissions at the study site increases during this time of day, which likely increases workload for both the individual intern and for the team as a whole. Alternatively, as the day progresses, the cumulative effects of fatigue may also play a role in subjective workload and ancillary resources available at the study site (ie, ultrasonography and some other radiology services) around this time. It seems likely that many or all of these factors contributed to the demonstrated increase in self-reported subjective workload.

Our finding that mean subjective workload is not correlated with the traditional measures of resident workload, such as patient census and number of admissions, is interesting. We hypothesize that these measures do

not fully capture the complexity and difficulty of the work done by residents.[25] It is also possible that mean workload is not the right way to quantify workload in an environment that changes as rapidly as the inpatient medicine wards. Perhaps the maximum self-reported subjective workload or the absolute variation in workload would be a better metric for representing workload over a given time period.

There are limitations to this study. First, it was a pilot study and included only 25 interns and was likely underpowered to detect certain differences or trends. Second, it was conducted before the 2011 duty hour rules went into effect, so the changes in workload seen during the middle of a 28-hour shift may be different with a maximum shift length of 16 hours. Third, this was conducted at a single site; therefore, the generalizability may be limited. Fourth, much of the data (ie, census and admission numbers) were self-reported, making reporting errors and recall bias possible. Finally, the Borg scale has not been validated in internal medicine settings. Instruments that have been validated in such settings[26,27] are longer and would not be feasible to give repeatedly during a shift without the risk of interrupting the subject's workflow.

During the past decade, hours on duty have decreased for residents in all specialties, yet many residents still suffer from burnout[28] and depression.[29,30] Workload is the next concept that should be explored as a contributor to these adverse outcomes. The amount of time spent on different tasks has changed over time, with a preponderance of effort put toward documentation tasks.[15,31,32] In addition, longer duty hours before the institution of the limits may have allowed for some decompression in residents' workload, and now that the hours are shorter, the gains from decreasing fatigue may be lost by work compression.

Conclusion

We measured workload over time for internal medicine interns during an overnight shift. Self-reported subjective workload may be a useful tool for identifying the contributors to resident workload. Understanding the contributors to workload would help residency program leaders better assign and allocate work; this, in turn, could lead to a safer environment for physicians and patients alike.

References

1 Philibert I, Friedmann P, Williams WT; ACGME Work Group on Resident Duty Hours. New requirements for resident duty hours. *JAMA*. 2002;288(9):1112–1114.

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

3 Landrigan CP, Rothschild JM, Cronin JW, Kaushal R, Burdick E, Katz JT, et al. Effect of reducing interns' work hours on serious medical errors in intensive care units. *N Engl J Med*. 2004;351(18):1838–1848.

4 Lockley SW, Cronin JW, Evans EE, Cade BE, Lee CJ, Landrigan CP, et al. Effect of reducing interns' weekly work hours on sleep and attentional failures. *N Engl J Med*. 2004;351(18):1829–1837.

5 Shanafelt TD, Bradley KA, Wipf JE, Back AL. Burnout and self-reported patient care in an internal medicine residency program [summary for patients in *Ann Intern Med*. 2002;136(5):I29]. *Ann Intern Med*. 2002;136(5):358–367.

6 Bready L. Chapter 13: the Graduate Medical Education Community's responsibility for producing a fully trained physician. In: Philibert IA, Amis S Jr; for the ACGME Task Force on Quality Care and Professionalism, eds. *The ACGME 2011 Duty Hour Standards: Enhancing Quality of Care, Supervision, and Resident Professional Development*. Chicago, IL: ACGME; 2011:81–86.

7 Philibert I. Education and patient care effects of resident workload restrictions: tackling a largely unexplored subject. *Mayo Clin Proc*. 2012;87(4):311–313.

8 Backs RW, Ryan AM, Wilson GF. Psychophysiological measures of workload during continuous manual performance. *Hum Factors*. 1994;36(3):514–531.

9 Jacobson CJ Jr, Bolon S, Elder N, Schroer B, Matthews G, Szaflarski JP, et al. Temporal and subjective work demands in office-based patient care: an exploration of the dimensions of physician work intensity. *Med Care*. 2011;49(1):52–58.

10 Wu AW, Folkman S, McPhee SJ, Lo B. Do house officers learn from their mistakes? [discussion in *Qual Saf Health Care*. 2003;12(3):227–228]. *Qual Saf Health Care*. 2003;12(3):221–226.

11 Fletcher KE, Parekh V, Halasyamani L, Kaufman SR, Schapira M, Ertl K, et al. The work hour rules and contributors to patient care mistakes: a focus group study with internal medicine residents. *J Hosp Med*. 2008;3(3):228–237.

12 Biaggi P, Peter S, Ulich E. Stressors, emotional exhaustion and aversion to patients in residents and chief residents—what can be done? [comment in *Swiss Med Wkly*. 2003;133(45–46):629]. *Swiss Med Wkly*. 2003;133(23–24):339–346.

13 Borg G. Simple rating methods for estimation of perceived exertion. In: Borg G, ed. *Physical Work and Effort*. Oxford, England: Pergamon Press; 1977:39–47.

14 Weinger MB, Slagle JS, Nwosu SK; V-A QWAT Research Team. Higher pre- and post-case workload ratings by anesthesiologists, surgeons, and OR nurses are associated with non-routine events [abstract]. *Proc Am Soc Anesthesiol Ann Mtg*. 2011.

15 Fletcher KE, Visotcky AM, Slagle JM, Tarima S, Weinger MB, Schapira MM. The composition of intern work while on call. *J Gen Intern Med*. 2012;27(11):1432–1437.

16 Borg G. Borg's range model and scales. *Int J Sport Psychol*. 2001;32:110–126.

17 Kurokawa T, Ueda T. Validity of ratings of perceived exertion as an index of exercise intensity in swimming training. *Ann Physiol Anthropol*. 1992;11(3):277–288.

18 Slagle JM, Weinger MB. Effects of intraoperative reading on vigilance and workload during anesthesia care in an academic medical center. *Anesthesiology*. 2009;110(2):275–283.

19 Cao CG, Weinger MB, Slagle J, Zhou C, Ou J, Gillin S, et al. Differences in day and night shift clinical performance in anesthesiology. *Hum Factors*. 2008;50(2):276–290.

20 Weinger MB, Reddy SB, Slagle JM. Multiple measures of anesthesia workload during teaching and nonteaching cases. *Anesth Analg*. 2004;98(5):1419–1425.

21 Weinger MB, Herndon OW, Gaba DM. The effect of electronic record keeping and transesophageal echocardiography on task distribution, workload, and vigilance during cardiac anesthesia. *Anesthesiology*. 1997;87(1):144–155.

22 Lamba R, Schapira MM, Singh S, Fletcher KE. Defining and measuring the effort needed for inpatient medicine work. *J Hosp Med*. 2012;7(5):426–430.

23 Weinger MB, Herndon OW, Zornow MH, Paulus MP, Gaba DM, Dallen LT. An objective methodology for task analysis and workload assessment in anesthesia providers. *Anesthesiology*. 1994;80(1):77–92.

24 Slagle J, Weinger MB, Dinh MT, Brumer VV, Williams K. Assessment of the intrarater and interrater reliability of an established clinical task analysis methodology. *Anesthesiology*. 2002;96(5):1129–1139.

25 Fletcher KE, Reed DA, Arora VM. Doing the dirty work: measuring and optimizing resident workload. *J Gen Intern Med*. 2011;26(1):8–9.

26 Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock PA, Meshkati N, eds. *Human Mental Workload*. Amsterdam, the Netherlands: North Holland Press; 1988.

27 Bertram DA, Opila DA, Brown JL, Gallagher SJ, Schifeling RW, Snow IS, et al. Measuring physician mental workload: reliability and validity assessment of a brief instrument. *Med Care*. 1992;30(2):95–104.

28 Ishak WW, Lederer S, Mandili C, Nikravesh R, Seligman L, Vasa M, et al. Burnout during residency training: a literature review. *J Grad Med Educ*. 2009;1(2):236–242.
29 Gopal R, Glasheen JJ, Miyoshi TJ, Prochazka AV. Burnout and internal medicine resident work-hour restrictions [see comment]. *Arch Intern Med*. 2005;165(22):2595–2600.
30 Landrigan CP, Fahrenkopf AM, Lewin D, Sharek PJ, Barger LK, Eisner M, et al. Effects of the Accreditation Council for Graduate Medical Education duty hour limits on sleep, work hours, and safety. *Pediatrics*. 2008;122(2):250–258.
31 Oxentenko AS, West CP, Popkave C, Weinberger SE, Kolars JC. Time spent on clinical documentation: a survey of internal medicine residents and program directors. *Arch Intern Med*. 2010;170(4):377–380.
32 Almromaihi D, Godfrey A, Dimoski T, Gunnels P, Scher E, Baker-Genaw K. Internal medicine residents' time study: paperwork versus patient care. *J Grad Med Educ*. 2011;3(4):550–553.