Estimating Minimum Program Volume Needed to Train Surgeons: When 4 × 15 Really Equals 90

Jennifer Chung, MD,[*] Andrea Obi, MD,[†] Ryan Chen,[‡] Wandi Lin, BSE,[‡] Siyuan Sun, BSE,[‡] Zixiao Chen,[‡] Anurag Gulati,[‡] Xun Xu,[‡] William Pozehl, BSE,[‡] F. Jacob Seagull, PhD,[*] Amy M. Cohn, PhD,[‡,§] Mark S. Daskin, PhD,[‡] and Rishindra M. Reddy, MD[*,†]

[*]University of Michigan Medical School, Ann Arbor, Michigan; [†]Department of Surgery, University of Michigan, Ann Arbor, Michigan; [‡]Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan; and [§]Center for Healthcare Engineering and Patient Safety, Ann Arbor, Michigan

INTRODUCTION: Work-hour restrictions have decreased flexibility in scheduling and reduced exposure to certain operative cases. These restrictions may affect a resident's ability to meet certification requirements, particularly for rare, unscheduled cases (e.g., cardiothoracic transplants). We developed a computer-based simulation model using variables such as case volume and program size to demonstrate the influence of these factors on the likelihood of certifying a set of residents on rare cases.

METHODS: We built a simulator to predict the probability of attaining certification for surgical residents, using cardiothoracic transplants as a test case. Inputs to the model included operating times, call schedules, and procurement travel times, as well as information on the distribution of times between transplants.

RESULTS: We simulated 100 years of schedules using our current system parameters of an average of 33 heart and 31 lung transplants per year, and assuming an Accreditation Council for Graduate Medical Education–compliant daily-rotating call schedule. Despite having enough transplants to certify all residents for lungs if all opportunities were distributed equally among residents, the certification rate achieved when constrained by arrival time (and call schedules) and work restrictions was only 55%. Our calculations show that meeting minimum transplant-certification requirements for all residents would require at least 1.5 times the expected number of annual transplants.

CONCLUSIONS: Our model enables analysis of a given program's ability to certify its residents based on program size and volume. These results could be used to design alternative scheduling paradigms to improve certification rates, without requiring reductions in certification requirements or program size. (J Surg 72:61-67. © 2014 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: surgery, graduate medical education, Scheduling, transplant

COMPETENCIES: Medical Knowledge, Practice-Based Learning and Improvement, Systems-Based Practice

INTRODUCTION

Ongoing restrictions on residency work hours implemented by the Accreditation Council for Graduate Medical Education (ACGME) in 2003 have forced residency and fellowship programs to implement new on-call schedules and have presented the challenge of meeting programs' training objectives within the reduced work hours. The feasibility of maintaining the same level of training, and ultimately competence, within the work-hour restrictions varies across the specialties that the ACGME regulates. There has been widespread concern within the surgical community regarding the effect of work-hour restrictions on the ability to train future surgeons and to achieve adequate case volumes necessary for competence.[1,2] The surgical literature so far has been inconclusive regarding the effect of restricted work hours on general surgery resident case volumes.[3,4] Connors et al.[5] specifically reviewed operative experience in cardiothoracic (CT) surgery and demonstrated declining case volumes after implementation of work-hour restrictions. Fairfax et al.[6] demonstrated that new ACGME-compliant

Correspondence: Inquiries to Rishindra M. Reddy, MD, Department of Surgery, University of Michigan, 1500 E. Medical Center Drive, Section of Thoracic Surgery, Ann Arbor, MI 48109; fax: (734) 615-2656; e-mail: reddyrm@med.umich.edu

1931-7204/$30.00 http://dx.doi.org/10.1016/j.jsurg.2014.07.010

call schedules resulted in reduced emergency case experiences for surgery residents.

Inflexible call schedules can result in different case volumes and experiences among residents, differences that may be magnified in rare, emergent cases. Before the institution of work-hour restrictions, residents had greater flexibility that allowed them to maximize their exposure to emergency surgeries. However, the work-hour restrictions have necessitated more rigid rotation and call schedules with limited flexibility, which may prevent residents from gaining exposure to emergency operations when not on call.[7] Because emergency cases are not scheduled, they have a high rate of variability and may demonstrate seasonality. For example, the occurrence of trauma operations is highly variable, but cases are rarer during cold, winter months. Due to this variability, it is unclear what minimum annual volume of each emergency case, and even nonemergent case, a training program needs to ensure all trainees acquire the minimum case numbers for board eligibility or certification.

CT transplants, which consist of heart and lung transplants, demonstrate the unpredictable nature of emergent cases and the potential for an uneven distribution of cases across residents operating under a rigid call schedule. Thus, they serve as a good proxy for the training challenges inherent in emergent surgical cases. Transplants are not required for board certification in surgery or thoracic surgery,[8] but are required for United Network of Organ Sharing certification to be an independently practicing transplant surgeon at a qualified transplant center.[9,10] We chose to evaluate CT transplants, as failure to achieve certification in transplantation will affect neither a trainee's board eligibility nor a training program's status, allowing us to evaluate our program and potentially others openly, without fear of residency review committee violations. Also, the analysis of CT transplants may have important long-term workforce implications. CT surgery overall has the oldest workforce of any surgical specialty,[11] with an estimated average age of 60 years (2011 Society of Thoracic Surgery Presidential Address). Given the potential decrease in resident case volumes and an aging workforce, it is important to understand the link between the random arrival of transplants, ACGME rules, and the ability to train and maintain the CT transplant workforce.

There are currently no tools available to analyze the effects of program size and operative volume on the probability of successfully graduating surgeons with minimum case requirements. This research, using CT transplants as a "test" case, provides a more rigorous method of analyzing, understanding, and predicting the rate of occurrence of emergency cases and the involvement of residents and fellows in these activities based on their call schedule. By using historical operative data, a computer simulation model will enable us to simulate the occurrence of emergent events in a given time period and thus calculate the number of cases performed by each resident in a fixed, ACGME-compliant call schedule. This approach has significant implications for all surgical fields with emergent operations, and the method demonstrated here could be extended to evaluate nonemergent cases as well.

METHODS

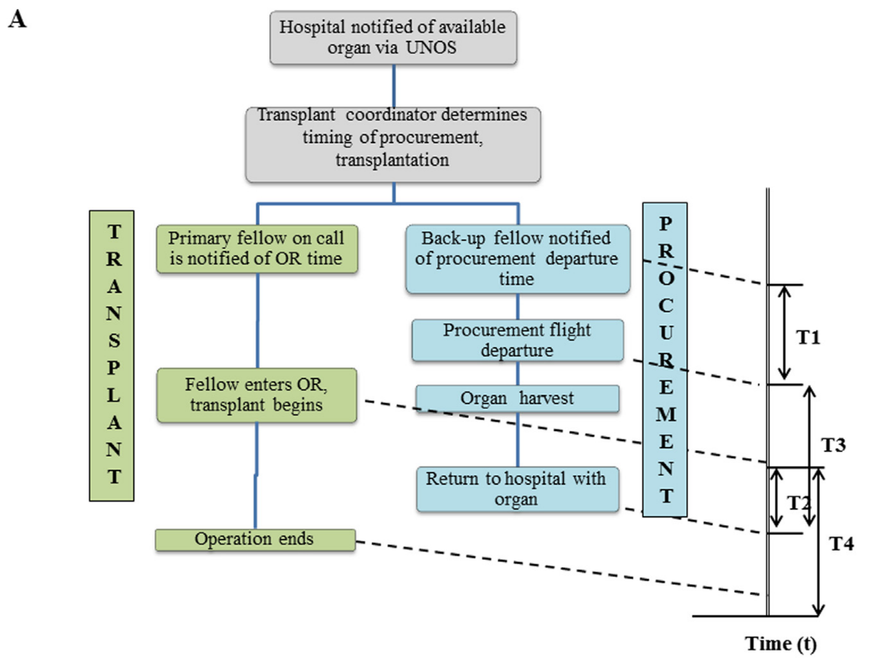
Institutional heart and lung operative case data were obtained from billing records and institutional transplant logs retrospectively. This study was approved by the institutional review board through a waiver of informed consent (HUM00054073).

This study used data from a 990-bed academic tertiary care center with a large volume of heart (8/130 in the United States in 2012 by volume) and lung (17/107) transplants.[12] Data were drawn from 36 months of lung and heart transplant operations (July 1, 2008 to June 30, 2011). The data gathered included all adults undergoing heart, single lung, and double lung transplants during this time period at our institution. Patients younger than 18 years were excluded. Parameters recorded included time and date of notification of organ availability, organ acceptance, departure and return from the procuring hospital, operating room (OR) entry, skin incision, skin closure, and OR exit, for each transplantation procedure.

Data Analysis

Critical events and the time between those events were identified. Critical events were defined as notification of organ availability, resident contact time, procurement departure time, procurement return time, OR entry time, and OR exit time for transplantation (Fig. 1A). These events outlined 4 key time intervals: the time between notification of a procurement and the flight departure (T1), the time between the start of the transplant surgery and the arrival of the organ (T2), the time between the departure and return of the procurement flight (T3), and the total transplant surgical time (T4). The data were then analyzed in the following 4 ways: calculation of the time intervals between critical events, analysis of the distribution of the time between sequential organ arrivals (organ interarrival times), construction of mathematical models to simulate the distribution and occurrence of future events, and addition of ACGME constraints based on the call schedule (Fig. 1B). The events that are initiated when a potential organ becomes available were broken down into key components in the work flow of both procurement and transplantation events.

Based on institutional historical data, the distributions of each of the key time intervals in the work flow of procurement/transplantation events were determined using mathematical equations that represented the stochastic nature of transplant events. A model was built to predict the number of transplant events that might occur within a given time period as well as the estimated duration of the various components of these transplants, providing the necessary data to predict the total hours of resident involvement.



B

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
6a-6p	**A/B/C/D**	**A/B/C/D**	**A/B/C/D**	**A/B/C/D**	**A/B/C/D**	**B**	**C**
6p-6a	**A**	**B**	**C**	**D**	**A**	**B**	**C**

FIGURE 1. (A) Critical events in transplant work flow. The occurrence of critical events and time in between events (T1-T4) in the transplant and procurement work-flow process was analyzed, and mathematical equations were developed to describe their distributions. (T1) is the time between notification of a procurement and the flight departure, (T2) the time between the start of the transplant surgery and the arrival of the organ, (T3) the time between the departure and return of the procurement flight, and (T4) the total transplant surgical time. (B) ACGME-compliant call schedule. This is a standard ACGME-compliant, 80-hour workweek call schedule with 4 residents (A-D) in rotation. The residents cover their normal assigned services on each weekday from 6 am to 6 pm. Night call, from 6 pm to 6 am, is taken from home on a rotating schedule with senior residents reporting to the hospital for any emergent situations. Weekends are split to give each resident on average 1 day off in 7. With this call schedule, there is a possible maximum of 86 hours worked if a resident is required to be in-house for 2 full on-call shifts, and a minimum of 60 hours worked if no call shift requires any in-house attention.

The data generated by the model were then overlaid on a standard ACGME-compliant resident call schedule in which each resident in the rotation is scheduled in the hospital for no more than 80 hours (Fig. 1B). Our call schedule has 4 people in the rotation, based on our institution's current CT program cohort of 2 residents per year in a 2-year program. Using a time period of 2 years (the length of our fellowship program), 50 repetitions of the simulator were run to generate a representative data set of 100 years of transplants. The predicted number of cases per year over multiple repetitions generated a model with a normal distribution, estimating 32.8 mean annual heart transplants and 31.3 mean annual lung transplants.

Analyses were then performed to investigate the effects of key independent variables on our outcomes of interest. Inputs included program transplant volume (n), the number of residents in the rotation or program size (r), and the number of cases needed to achieve certification (m). The outcomes of interest were the individual certification rate (I), the probability that any given resident will achieve certification, and the program certification rate (P), the probability that all residents in a program will meet the case requirements for certification. In our analyses, the scheduling period (t) was set at 720 days, the program size was set at 4 residents ($r = 4$) and the program requirements for certification were set at 15 procedures ($m = 15$) for lung transplants and 20 ($m = 20$) for heart transplants, based on United Network of Organ Sharing certification criteria. The specific mathematics used for the modeling is presented in the Supplementary material.

RESULTS

Effects of Program Volume and Program Size on Individual and Program Certification Rate

The number of emergent cases, which were heart or lung transplants or both in our model, follows a Poisson distribution with an arrival rate of 1 in 11.35 days, and an average 2-year total of 64 cases. Given the similar arrival rates for both hearts and lungs, the main distinction between the model inputs for the 2 scenarios were the case requirements, set at 20 for hearts and 15 for lungs.

Analyses of transplant program volume revealed a nonlinear relationship between program volume (n) and the probability that any given resident will achieve certification (I) (Fig. 2A).

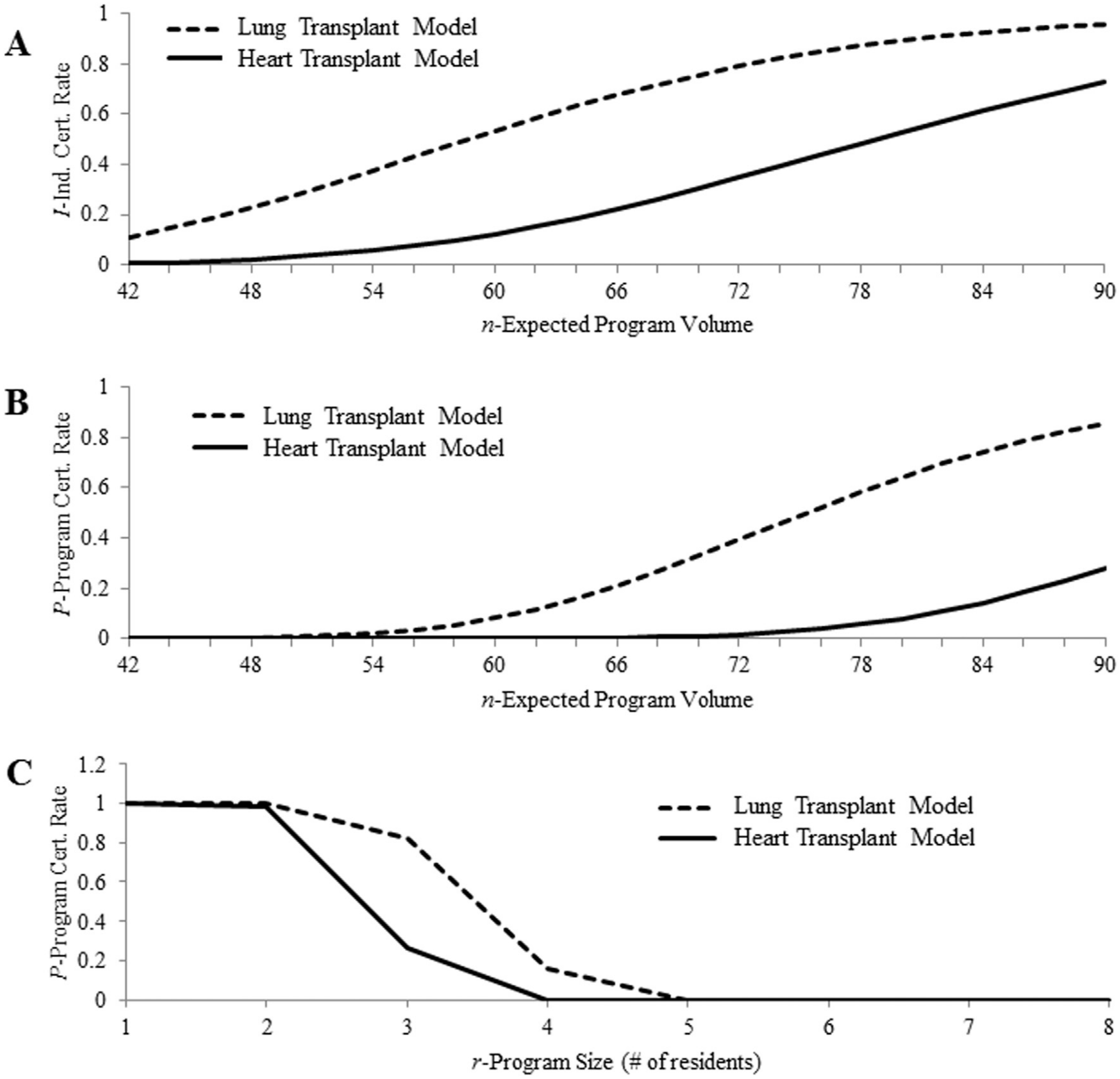


FIGURE 2. (A) Individual certification rates will rise more quickly for lung transplants as compared with heart transplants, given the lower requirements to achieve certification. Even at 60 transplants per year, the minimum number needed to train all 4 trainees with 15 transplants each, the individual certification rate (*I*) is just more than 50%, indicating that on an average only 2 of 4 trainees would have achieved certification status owing to the random arrival of transplants. (B) Program certification rates (*P*) reflect similar limitations. At 60 transplants, the likelihood that all 4 trainees are certified is just under 10%. Almost 90 transplants per year would be needed to ensure that all 4 trainees achieve certification status (i.e., perform a minimum of 15 transplants each). Program certification rates will rise more slowly than individual rates with an increase in volume. (C) Program certification rates (*P*) for operations with higher minimum case requirements (heart transplants, $m = 20$) are more sensitive to changes in program size than those with lower case requirements (lung transplants, $m = 15$).

The program volume necessary for this certification rate to exceed 90% is substantially greater than just the product of the number of residents (r) and the minimum number of cases required for certification (m). For example, in a program with 4 residents, each of whom needs 15 lung transplants to qualify for certification, the probability that any individual resident achieves certification is only 55% when the program volume is 60 cases over the 2-year period, or averages 30 cases per year. The graph demonstrates that at least 90 lung transplant cases would be necessary for I to approach 1, or 100%. A similar curve is seen with heart transplants, in which even greater numbers are needed to achieve the same rates of certification (I) given the higher case requirements ($m = 20$); to achieve the same certification rate of approximately 55%, the required program volume for heart transplants rises to 80.

There is also a nonlinear relationship between program volume (n) and the probability that all residents will achieve certification (P) (Fig. 2B). When the average lung transplant volume is 60 over 2 years, at which point theoretically there are adequate cases for each of the 4 residents to achieve certification, there is just less than a 10% chance that all trainees will reach their target case load. The graphs indicate that there is a threshold effect, a lower limit of program volume that must be exceeded for there to be a nontrivial probability of all residents being certified. For lung transplants, this lower limit is 56. Significantly larger program volumes are

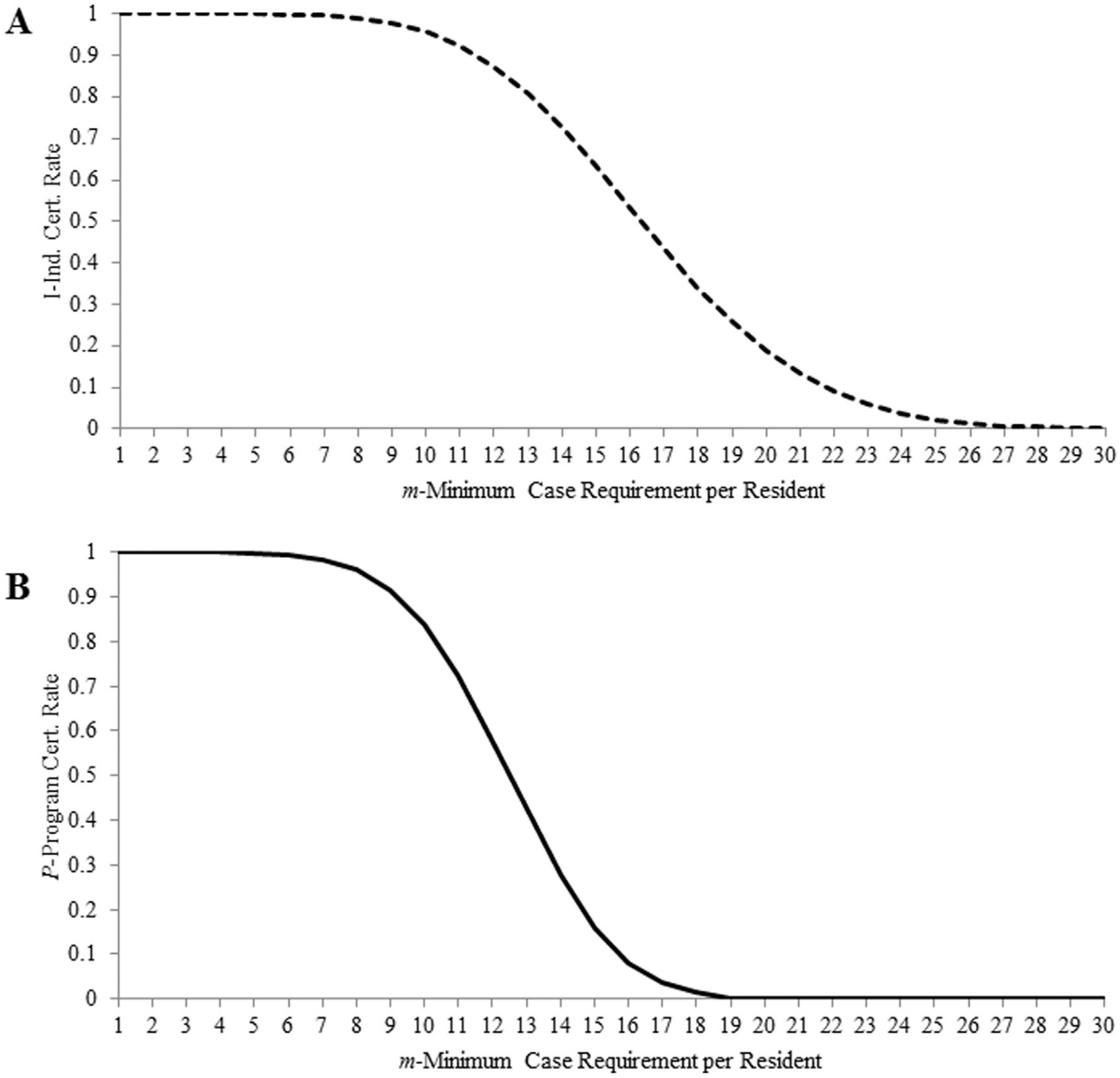


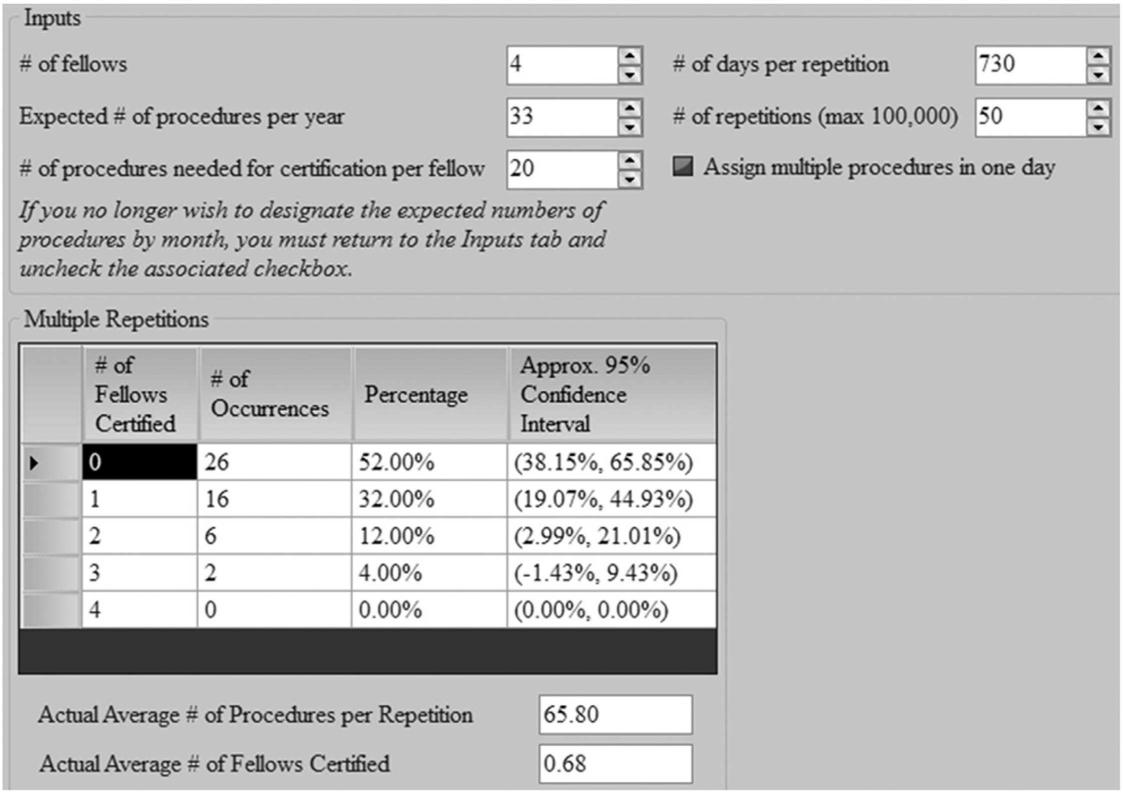
FIGURE 3. (A) The variation in individual certification rates (*I*) as a function of increasing minimum resident requirements (*m*) was investigated at a fixed volume of 64 cases per 2 years and program size of 4 residents. A decrease in certification rates is seen when programs require more than 8 cases. (B) Using the same fixed variables as in (A), the relationship between program certification rates (*P*) and minimum requirements (*m*) demonstrates a similar and more precipitous decrease when the program requirement is greater than 10.

needed to approach a 100% probability of certifying all residents ($P = 1$). From a different perspective, we can also consider the effect of program size and the number of trainees in the program (r), on certification rates. When the program size increased beyond 2 residents, there was a pronounced decrease in P, especially in heart transplants (Fig. 2C).

Effect of the Case Requirement Number on Individual and Program Certification Rate

Similarly, when the minimum case requirements (m) are changed, but the total number of cases available (n) is constant, there is a critical point after which incremental changes result in drastic reductions in certification rates. This critical point for individual residents is when the required number exceeds 11 (Fig. 3A), and for the program as a whole when the required number exceeds 10 (Fig. 3B). As case requirements increase, a greater reduction is seen in program certification rates compared with individual certification rates. This is an expected result as program certification requires that all residents achieve certification, as opposed to the likelihood of certification for an individual resident.

These key relationships between program volume, size, and requirements on outcomes form the basis of our simulator. The simulator interface (Fig. 4) allows the user to specify the key variables of program size (“no. of fellows,” f), program volume (“expected no. of procedures per year,” n), and program requirements (“no. of procedures needed for certification per fellow,” t). The scheduling period (“no. of days per repetition,” d) and the number of repetitions of the simulation are also adjustable inputs. The model assumes that all residents follow an ACGME-compliant call schedule as


Inputs

\# of fellows: 4

Expected # of procedures per year: 33

\# of procedures needed for certification per fellow: 20

\# of days per repetition: 730

\# of repetitions (max 100,000): 50

☑ Assign multiple procedures in one day

If you no longer wish to designate the expected numbers of procedures by month, you must return to the Inputs tab and uncheck the associated checkbox.

Multiple Repetitions

# of Fellows Certified	# of Occurrences	Percentage	Approx. 95% Confidence Interval
0	26	52.00%	(38.15%, 65.85%)
1	16	32.00%	(19.07%, 44.93%)
2	6	12.00%	(2.99%, 21.01%)
3	2	4.00%	(-1.43%, 9.43%)
4	0	0.00%	(0.00%, 0.00%)

Actual Average # of Procedures per Repetition: 65.80

Actual Average # of Fellows Certified: 0.68

FIGURE 4. The simulator interface allows the user to adjust for several variables. In addition, the simulation can be run multiple times, up to a maximum of 100,000, to generate large data sets and eliminate bias. As an example, in a scenario where there is an average of 33 heart transplants performed each year in a program with 4 fellows, each needing 20 heart transplants by the end of a 2-year training program, the simulator shows that 32% of the time (CI: 19.07%-44.93%) no one gets certified and the average number of fellows certified is less than 1.

outlined in Fig. 1B. The simulator then calculates the chance of achieving certification, or meeting the case requirements.

Validation

The model was verified by temporarily replacing the stochastic variables, such as the procurement and transplant times, with constants. This allowed us to predict the outcomes of the model and identify any suspicious outliers. This method identified some logical flaws in the model and typos in the code, which were corrected. The fixed variables were returned to stochastic variables and the model output was reviewed for any unexpected values.

The model was then tested against historical institutional transplant data from 2010, which was set as the control data set. Each repetition of the model yielded a mean of approximately 33 for hearts and 31 for lungs, which corresponded to the mean of the control data set. The standard deviation of 6 is acceptable in comparison with that of the mean.

DISCUSSION

The analysis underlying the development of our simulator represents the first published attempt to formally assess the effects of a stochastic number of emergent cases along with various real-life constraints on the potential training outcomes of surgical residents using a mathematical modeling system. Our results demonstrate that under a rigid call schedule designed to comply with ACGME work-hour restrictions, residents are unlikely to meet training requirements despite the perception of adequate total program volume and resident participation in every transplant. Our institutional data reflect even lower rates of transplant certification after 2003, than our model would suggest, given our training program's efforts to comply with work hours and the use of surgical assistants rather than CT residents for many transplants. Our analyses have shown that the program volume must be significantly higher than might have been previously thought. Simply multiplying the number of trainees (4) by the number of cases each trainee needs (15 for lung transplants, or 20 for heart transplants, yielding totals of 60 and 80, respectively) resulted in very low individual certification rates (55%) and even lower program certification rates (10%). The actual program volume needed for all trainees to meet the minimum case requirements with high probability was more than 1.5 times higher than the product of the number of trainees and the minimum case numbers. The framework developed by our team provides a method of estimating the minimum volume needed for achievement of high program throughputs and can be used as a tool for early recognition of volume deficits, allowing program directors to make adjustments as needed for projected shortfalls. The simulator interface allows for the adjustment of independent variables to reflect the various programs' specific characteristics and predicts the resulting likelihood of trainee case-volume achievement.

Our model currently deals with the specific condition of emergent surgical cases based on historical transplant data, but we hypothesize that this conceptual framework could be further expanded to apply to nontransplant emergent cases and even elective/scheduled cases. Future work includes analyzing other CT program transplant case data with alternative call schedules and different resident numbers. We will also analyze nontransplant operations, both emergent and elective. A future application includes using the simulator to identify ongoing adjustments in call schedules and training assignments that are needed when it becomes clear that a trainee is at risk of failing to meet certain case requirements, and to help direct educational programming and improve training opportunities. We can also develop alternate schedules with greater ACGME flexibility to maximize training opportunities and allow certain residents interested in transplants to maximize their transplant exposure, if desired.

Limitations

Our study data were limited to a single institution's historical data for heart and lung transplantations and a single call schedule. Although we believe that our model would be generalizable to other transplant programs and to most emergent-type cases, further validation of this hypothesis is necessary.

REFERENCES

1. Antiel RM, Thompson SM, Hafferty FW, et al. Duty hour recommendations and implications for meeting the ACGME core competencies: views of residency directors. *Mayo Clin Proc*. 2011;86(3):185-191.
2. Antiel RM, Van Arendonk KJ, Reed DA, et al. Surgical training, duty-hour restrictions, and implications for meeting the Accreditation Council for Graduate Medical Education core competencies: views of surgical interns compared with program directors. *Arch Surg*. 2012;147(6):536-541.
3. Bruce PJ, Helmer SD, Osland JS, Ammar AD. Operative volume in the new era: a comparison of resident operative volume before and after implementation of 80-hour work week restrictions. *J Surg Educ*. 2010;67(6):412-416.
4. Christmas AB, Brintzenhoff RA, Sing RF, et al. Resident work hour restrictions impact chief resident operative experience. *Am Surg*. 2009;75(11):1065-1068.
5. Connors RC, Doty JR, Bull DA, May HT, Fullerton DA, Robbins RC. Effect of work-hour restriction on operative experience in cardiothoracic surgical residency training. *J Thorac Cardiovasc Surg*. 2009;137(3):710-713.
6. Fairfax LM, Christmas AB, Green JM, Miles WS, Sing RF. Operative experience in the era of duty hour restrictions: is broad-based general surgery training coming to an end? *Am Surg*. 2010;76(6):578-582.
7. Feanney MA, Scott BG, Mattox KL, Hirschberg A. Impact of the 80-hour work week on resident emergency operative experience. *Am J Surg*. 2005;190(6):968-972.
8. American Board of Thoracic Surgery Index Case Requirements for Residents. Available at: 〈https://www.abts.org/root/home.aspx〉 Accessed 28.04.13.
9. Membership and Personnel Requirements for Heart Transplant Programs. Available at: 〈http://optn.transplant.hrsa.gov/policiesAndBylaws/bylaws.asp〉 Revised 18.04.13.
10. Membership and Personnel Requirements for Lung Transplant Programs. Available at: 〈http://optn.transplant.hrsa.gov/policiesAndBylaws/bylaws.asp〉 Revised 18.04.13.
11. American College of Surgeons Health Policy Research Institute. The Surgical Workforce in the United States: Profile and Recent Trends [Internet]. Available at: 〈http://www.acshpri.org/documents/ACSHPRI_Surgical_Workforce_in_US_apr2010.pdf〉; 2010 Updated April 2010.
12. Data Reports from Organ Procurement and Transplantation Network. Available at: 〈http://optn.transplant.hrsa.gov/data/〉 Accessed 28.04.13.

SUPPORTING INFORMATION

Supplementary material cited in this article is available online at doi:10.1016/j.jsurg.2014.07.010.