Proceedings of the 2022 Winter Simulation Conference
B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C.G. Corlu, L.H. Lee, E.P. Chew, T. Roeder, and P. Lendermann, eds.

USING DISCRETE-EVENT SIMULATION TO ANALYZE THE IMPACT OF VARIATION ON SURGICAL TRAINING PROGRAMS

Fumiya Abe-Nornes
Samir Agarwala
Nathan Smith
Rachel Zhang
Amy Cohn

Industrial and Operations Engineering
University of Michigan
1205 Beal Avenue
Ann Arbor, MI 48109, USA

Angela Thelen

Center for Healthcare Outcomes & Policy
University of Michigan
2800 Plymouth Road
Ann Arbor, MI 48109, USA

Rishindra Reddy

Thoracic Surgery, University Hospital
University of Michigan
1500 E Medical Center Dr
Ann Arbor, MI 48109, USA

Brian George

General Surgery, University Hospital
University of Michigan
1500 E Medical Center Dr
Ann Arbor, MI 48109, USA

ABSTRACT

In this paper, we use discrete-event simulation in an attempt to highlight the consequences of variability in surgical training. Under the current training model, case volume minimums are being used as a surrogate measure of a surgical trainee's competency for a given operation. However, this assumes that 1) learning is a binary measure, 2) there is no variability in training opportunities, and 3) all trainees learn at the same speed. Our model addresses these variables by allowing the user to manipulate the distribution of continuous learning curves and arrival rates, simulating the competency outcomes of a surgical training model. The results demonstrate that when increasing the variability in learning speeds or decreasing the training opportunities, competency outcomes for common procedures such as appendectomies remain relatively unaffected. However, for rarer procedures like mediastinoscopies, these variabilities result in a greater proportion of decreasingly competent trainees, potentially endangering patient safety.

1 INTRODUCTION

1.1 Background

The U.S. educational model for training surgeons is largely experience-based. The intent is that surgical trainees perform a number of procedures during their residency under progressively decreasing supervision until competency is achieved and they are fully prepared for independent practice. The way this is currently operationalized is that trainees are expected to perform a predetermined number of each procedure type within a predetermined window of time (Accreditation Council for Graduate Medical Education, 2019). Recent studies raise concerns that this approach does not ensure adequate training for all learners (George et al., 2017). In particular, the approach does not recognize variability such as the speed in which trainees acquire competency and the number of surgical opportunities that may arise over different training periods.

In the U.S., most surgical residency training programs last at least five years, depending on the specialization. Throughout this time, residents will rotate through different subspecialties and services like cardiology or emergency medicine on most often a monthly basis (Accreditation Council for Graduate Medical Education, 2022). On each service, they will observe and execute different procedures. Figure 1 shows an example of how general surgery rotations repeat each year but with potentially different timelines. The Accreditation Council for Graduate Medical Education (ACGME) considers a resident to be competent in a particular surgical technique if they complete a minimum number of relevant medical procedures. Different categories of surgery require a different minimum number of cases. For example, residents are required to complete five thoracotomies versus eighty five biliary surgeries (Accreditation Council for Graduate Medical Education, 2019).

When residents begin their training in a residency program, they are assigned a fixed schedule that outlines the different surgical services they are on and for how long. As an example, a resident may spend their first month on General and Vascular surgery, moving to Transplant surgery their second month, and then rotating to Endocrine surgery their third month. It is also common to be on the same service multiple times throughout the program, such as being on the Plastic surgery service for a month during the second year and then again for a month during the fourth year, or Acute Care surgery for a total of 6 times over the duration of the residency.

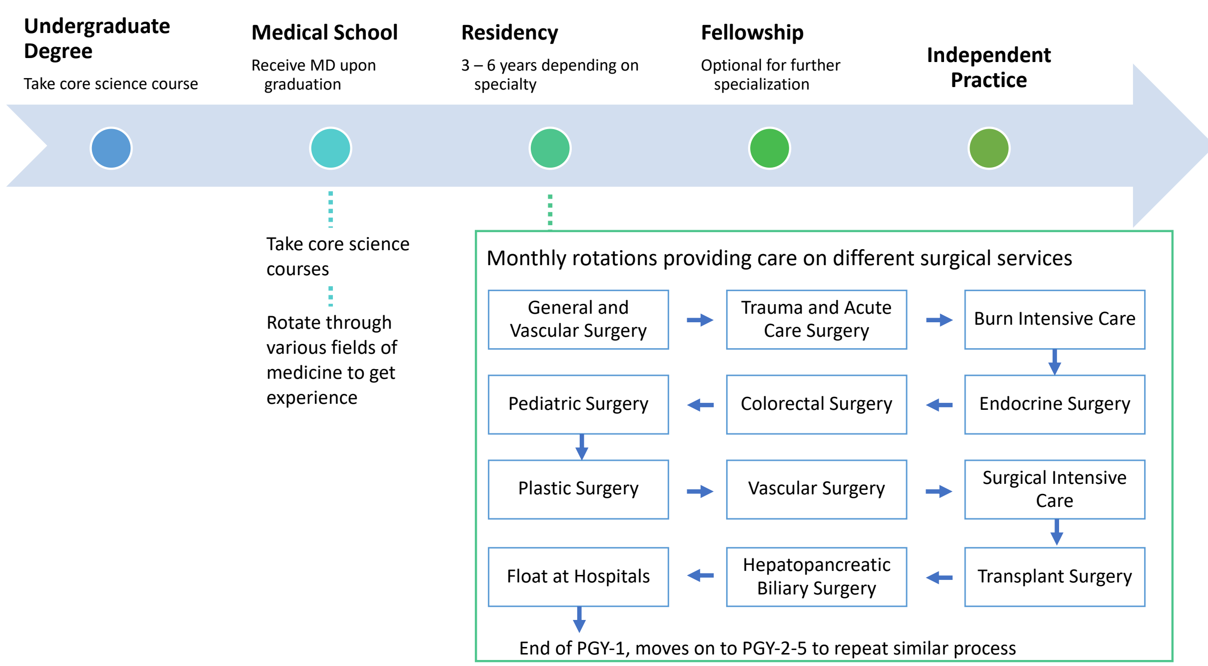


Figure 1: The path to becoming a surgeon in the United States.

ACGME outlines specific guidelines across the continuum of medical education with the goal of ensuring that residents and physicians are delivering safe, effective, and professional care. The case requirements for general surgery residents are static and not impacted by real-world variation in case arrival and individual resident learning curves. This model implies that all individuals learn at the same rate, that residents all have the same opportunity to perform the same number of cases, and that competency is a binary variable. These assumptions that the instructional model makes are broad generalizations that may allow a class of general surgery residents to graduate with a wider distribution of competency in certain procedures than ACGME would desire.

In this paper, we use discrete event simulation (DES) to model the impact of variability (both in procedure availability and in skill acquisition) on the ability of a residency program to ensure the development of adequate competence for all trainees. This research is based on a collaboration between the University of Michigan (UM) Department of Thoracic Surgery and the UM Center for Healthcare Engineering and Patient Safety (CHEPS). The goal was to identify opportunities for improvement of the surgical residency program by leveraging engineering techniques. In particular, recognizing the impact of variability is critical to the overall improvement of training programs.

1.2 Problem Definition

There are three key assumptions underlying the ACGME competency model that may detract from the goal of achieving real competency for every surgical resident. Specifically, assuming that 1) learning is a binary measure 2) there is no variability in the availability of surgical training opportunities, and 3) that all trainees learn at the same rate can suggest that a training model which is strictly time-based will be effective. In actuality, such a training model that doesn't consider the real world variability in training speed and procedure opportunity may fail to enable all residents to attain adequate competency.

In order to move away from the idea that competency is a binary measure, it is important to consider learning the skills necessary to become competent in a procedure to be continuous. When a trainee begins practicing a procedure, their progress in developing competency with each additional training opportunity they have will increase in a continuous, non-linear manner (Subramonian, 2004). This can be plotted as a function that will be referred to as a learning curve which can be seen in Figure 2. As they complete more training, their competency in that particular procedure will increase until eventually reaching an asymptotic level. In Figure 3, it can be seen that each additional case completed does not contribute a relatively significant amount of competency.

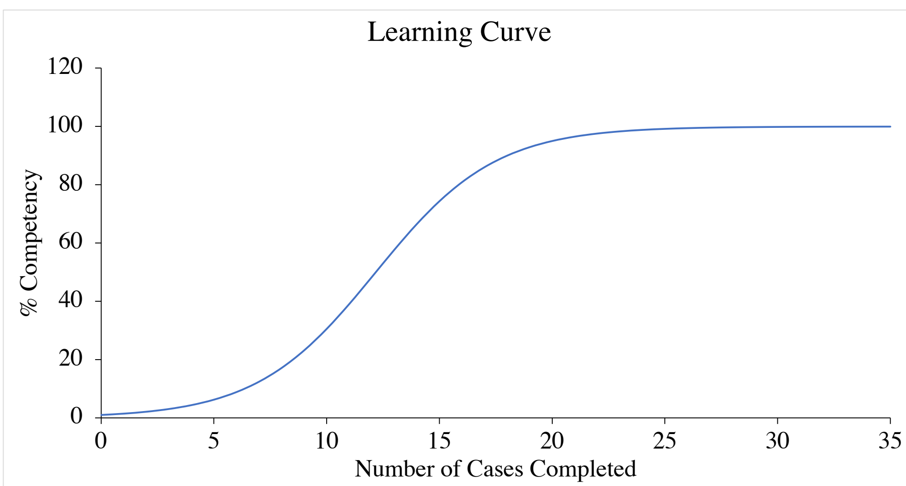


Figure 2: A sample learning curve for a given resident where they achieve 90% competency when they have completed the required 18 cases for that procedure. As the training progresses, the resident will eventually reach an asymptotic level, where each additional case completed does not significantly contribute to their overall competency in that procedure.

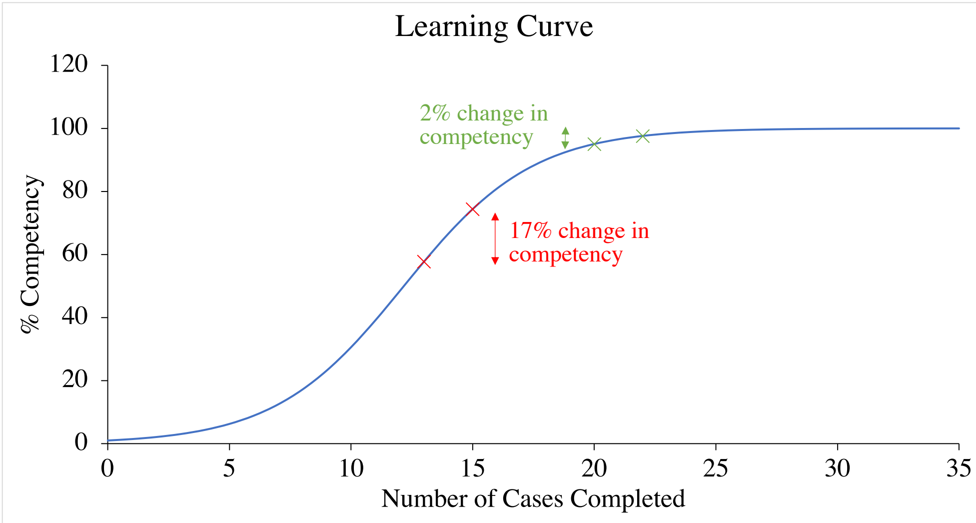


Figure 3: A sample learning curve for a given resident where they achieve 90% competency when they have completed the required 18 cases for that procedure. Over the course of the residency program, the variation in training opportunities can have a significant impact on the competency of that trainee in a procedure. For example, a trainee that completes 13 cases and a trainee that completes 15 cases of that same procedure could have a 17% difference in competency. However, when residents are towards the end of their learning curve, the difference in competency acquired is relatively smaller with each additional case completed. For example, a resident who completes 20 cases will only have a 2% difference in competency from a resident who completes 22 cases of the same procedure.

Procedure opportunity can vary based on the program, where large general surgery residency programs associated with prominent hospitals have enough inflow of different cases to maximize the possibility that each resident will meet the minimum case requirements. However, the real life variability in the types of cases that arrive in a given month to smaller programs could make it difficult for all residents to meet case minimums. In light of recent circumstances surrounding the COVID-19 pandemic, anomalies that slow/shut down the American healthcare system must also be factored into the potential variability in cases. Elective surgeries were completely shut down throughout the United States amidst the pandemic and many residents were asked to help treat COVID-19 patients instead of continuing their curriculum (Aziz et al. 2021).

Some procedures like hernia repair are very common, and residents likely have no problem fulfilling their case requirements. Other procedures however, like mastectomies, are more rare. For example, a given hospital with a residency program receiving on average 12 mastectomies per year would average to be 1 procedure per month. Due to the nature of a fixed schedule, if there were 12 residents in that program that were each on the associated service for one month, although some trainees may receive exactly one mastectomy while on the service, others may receive 2 or 0. Although the average number of mastectomies completed across the trainees would be 1 per resident, the variation and unpredictability in procedure availability demonstrates how it could impact the consistency of training opportunities. Ultimately, this could result in varying levels of competency across the residents.

Furthermore, each general surgery resident learns new skills at a different rate with each new chance to perform a procedure. Some residents may reach competency before completing the minimum number of cases. More importantly, there is another population of residents that learn skills at a slower rate who may be incorrectly deemed competent after meeting ACGME’s minimum case requirements, despite lacking the real competency level that would ensure safe practice. In Figure 4, the variation in learning curves arriving from different rates of learning is represented.

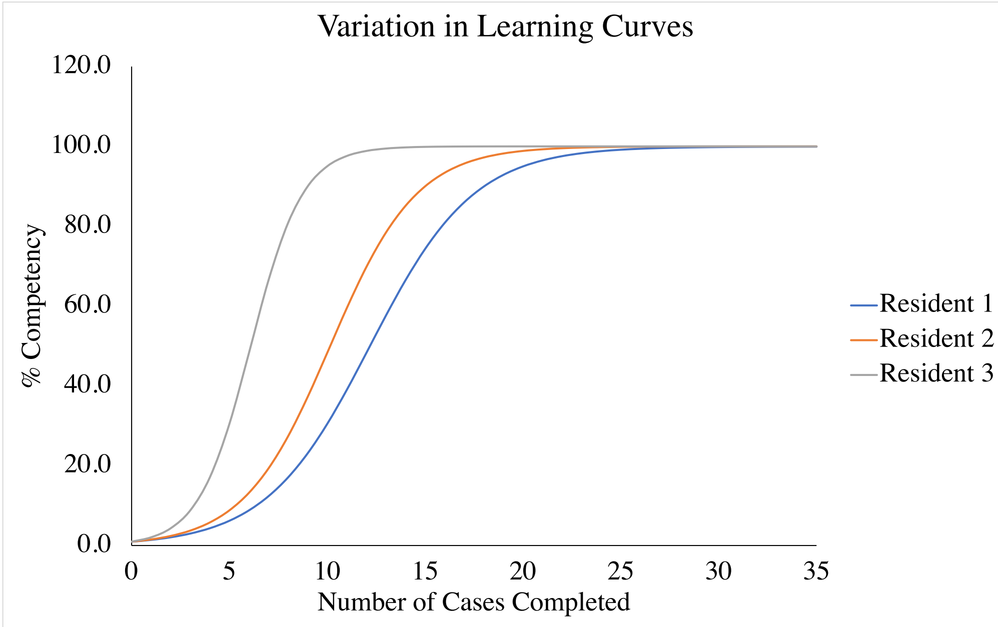


Figure 4: Three different learning curves for three different residents for a given procedure. ACGME case numbers imply that all trainees learn at the same rate. This figure displays how for a given procedure, different trainees may require different numbers of cases completed in order to reach the same level of competency. As a result, it is critical to assess the potential impact the variation of learning curves may have on residency program outcomes.

In this paper, we present a DES model to evaluate the impact of these underlying assumptions, so as to motivate the development and analysis of more flexible, trainee-focused alternatives. By building a simulation model and running computational experiments where we changed the level of variability, competency outcomes for different surgical procedures were assessed.

2 SIMULATION MODEL

Our model was built using C++, where the inputs to the simulation were parameters for distributions that would determine the number of cases a resident would receive for a given procedure and how many cases the resident must complete in order to become 90% competent in that procedure (also known as the learning curve parameter, LCP). Using the LCP, a learning curve for each trainee for each procedure is generated, where their competency in that procedure is based on the number of cases they completed over the duration of their residency program. We applied this model to compute the total competency gained by residents in various procedures after rotating through a general residency program of block-scheduled training.

The user of the simulation model selects the procedures within a service category and assigns each resident to one service per rotation. For example, only a resident assigned to the Thoracic service in July will be able to perform thoracotomies in July.

In order to simulate the number of cases of a given procedure a resident gets exposed to, the user is able to select a distribution and its associated parameters to model arrival rates. As an example, if a user were to input the parameters for the distribution of the arrivals of colectomies, every time the trainee is on the associated service, the model will randomly draw a value from that distribution that will serve as the number of training opportunities for that procedure on that rotation. As residents rotate through their schedule, sometimes being on the same service multiple times, the total number of cases completed for each procedure will be stored. These values will then be used to calculate competency through the use of their learning curve, which can be seen in Appendix A.

As there is also variation in the learning speeds of different trainees, the user will be able to define a distribution and its associated parameters to represent the possible range of learning curves for the residents. When the simulation is run, the resident will be assigned a LCP for every procedure, which is then used to create the learning curve for each procedure.

In summary, when the simulation is run, the code will generate a sample trainee, with the number of arrivals for each procedure drawn from their associated distributions based on the resident's schedule, as well as the trainee's LCP from their associated distributions for the same procedures. Once the simulation has rotated through the entirety of the resident's schedule, the total number of cases completed for each procedure is used to calculate the competency of the resident in each of the procedures using the associated learning curves. This process is repeated for a user-defined number of replications, where at the end, summary metrics such as the mean and median competency and the distribution of competencies for all residents are produced. A flowchart that demonstrates the process of one replication of the simulation model can be seen in Figure 5.

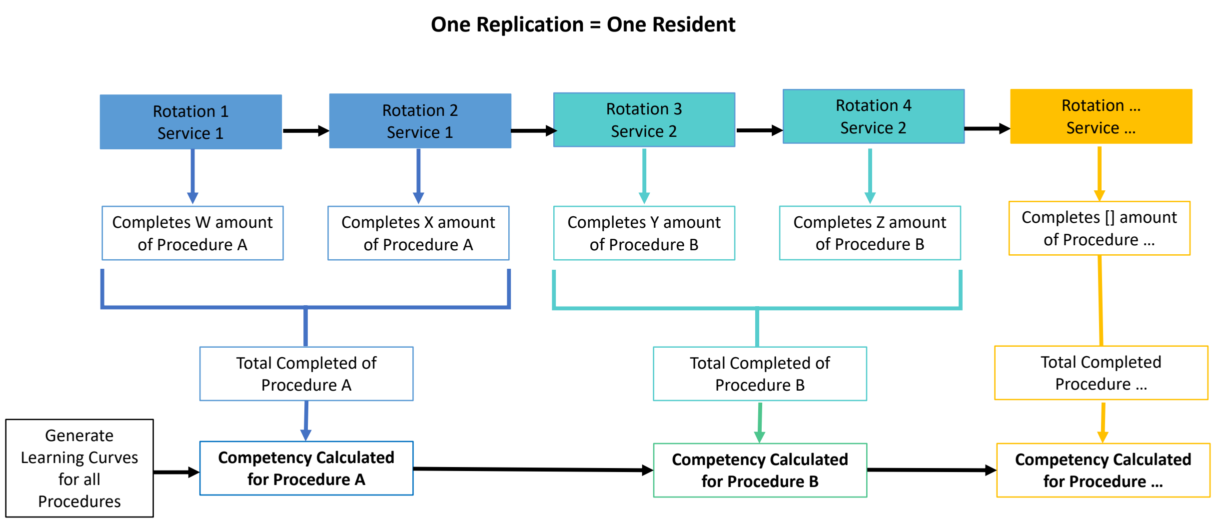


Figure 5: An overview of one replication of our simulation model. This process will repeat for the user-defined number of replications, where each replication can be thought of as one resident going through the residency program.

3 COMPUTATIONAL EXPERIMENTS AND RESULTS

3.1 Simulation Inputs

The following computational experiments are based on real-world parameters from a general surgery residency program in a major academic U.S. medical center. To compare and contrast the impact of variation on a common procedure versus an uncommon procedure, the competency outcomes for appendectomy and mediastinoscopy were investigated. The goal was to assess the potential impact that variation in arrivals and learning curves could have on competency outcomes. In this first scenario, we simulate when there is no variation in procedure opportunities nor learning speeds, followed by a second scenario with only variation in procedure opportunities. Finally, a third scenario combines variation in learning curves with variation in procedure opportunities.

The fixed schedule assigns the trainee to 12 months on the acute care service and 2 months on the thoracic service over the duration of their training. These are the specialties where a resident would receive their training opportunities for appendectomies and mediastinoscopies, respectively.

The ACGME case requirements of 40 and 5 for appendectomies and mediastinoscopies, respectively, were used as the parameter mapping to 90% competency on the baseline deterministic learning curve, which was done because the learning curve is generated from the LCP which defines the number of cases required to achieve 90% competency (see Appendix A). The values of 40 and 5 were also used as the mean for the Normal distribution used to generate learning curves to demonstrate the variation among residents (Accreditation Council for Graduate Medical Education, 2019; Accreditation Council for Graduate Medical Education, 2017). Standard deviations for the learning curve distributions were increased across experiments to assess the impact this increase in variation had on competency outcomes.

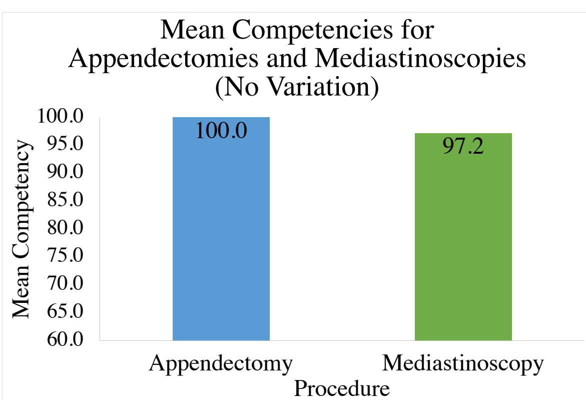
Finally, we used a Poisson distribution to represent independently-occurring procedural opportunities in a fixed time interval, with means of 38 and 3 per month, respectively, for appendectomies and mediastinoscopies, based on historical data (Gart, 1975). Table 1 shows a summary of our inputs in the simulation model. A total of 10,000 replications were run for each scenario.

Table 1: A table showing a summary of the inputs to the model for the computational experiments.

	Associated Service, Number of months on Service	Arrival Rate Distribution	Arrival Rate Distribution Parameters	Learning Curve Distribution	Learning Curve Distribution Parameters
Appendectomy	Acute Care, 12	Poisson	Mean = 38	Normal	Mean = 40 S.D. = Varied
Mediastinoscopy	Thoracic, 2	Poisson	Mean = 3	Normal	Mean = 5 S.D. = Varied

3.2 Scenario 0

We begin by presenting a baseline scenario where all residents receive equal opportunities for both procedures, and they learn at the same rate; the simulation model allows us to easily compute the outcomes for this simple case.



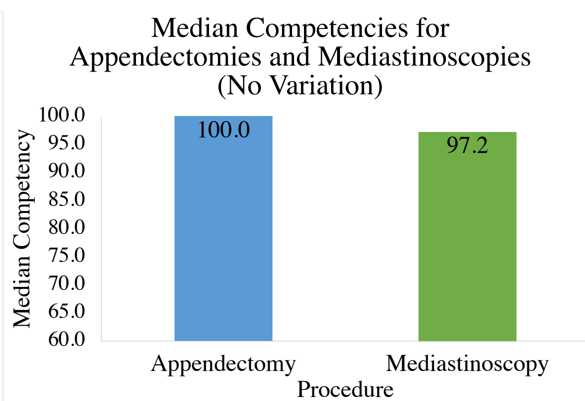


Figure 6A and 6B: Mean and median competencies for appendectomy and mediastinoscopy when no variation is introduced.

With this control scenario, there was no variation in competency observed. This makes sense, as everyone has equal training opportunities and has the same learning curve. Every simulated resident should be getting 38 appendectomies per month for 12 months and 3 mediastinoscopies per month for 2 months, which safely guarantees that everyone surpasses the required 40 and 5, respectively, achieving at least the 90% target goal. Given that there is no variability, the median and mean competency are of course equal. This scenario demonstrates how the current ACGME standards assume residential trainees develop competency.

3.3 Scenario 1

Next, we assess the impact of varying arrival rates by introducing a Poisson distribution with a mean of 38 and 3 for the arrival rate of appendectomies and mediastinoscopies, respectively, while keeping the learning curves consistent for all trainees.

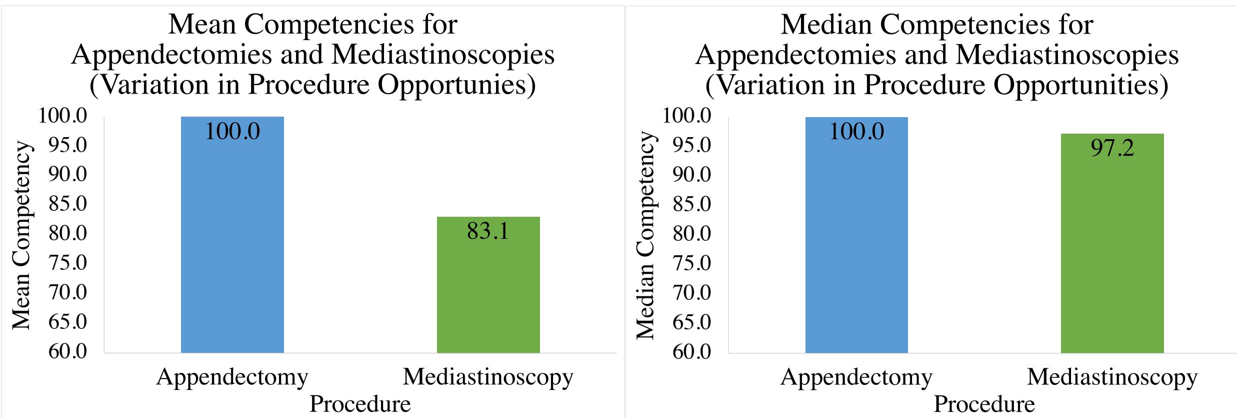


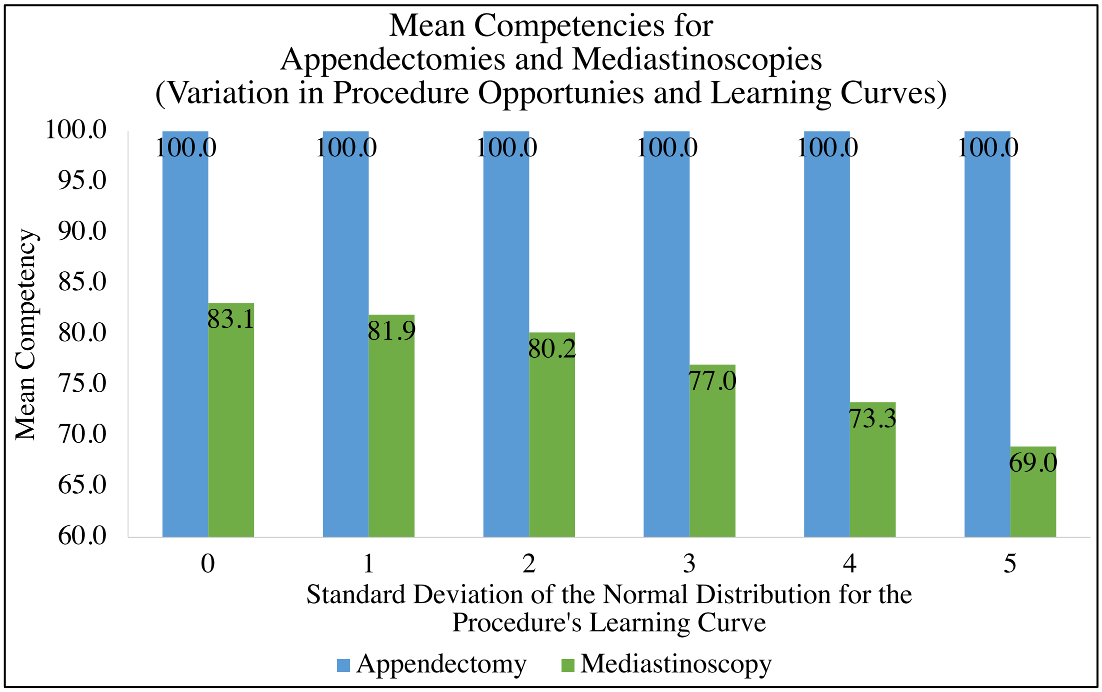
Figure 7A and 7B: Mean and median competencies for appendectomy and mediastinoscopy when variation in training opportunities was introduced.

Results of this simulation can be seen in Figure 7A and 7B, where it can be seen that varying the arrival rates for appendectomies did not greatly affect the competency outcome because there is an abundance of these procedures and the arrival rate over the residency program is well beyond the required number of 40 cases to achieve 90% competency on the learning curve. There is ample time for all residents to receive training opportunities, and we observe all residents still achieving full competency.

With the Thoracic service, however, there are relatively few training opportunities, and introducing variation in arrival rates substantially impacts the competency outcome. The median competency score of 97.2% may well be viewed as a success by program directors with limited experience exploring stochasticity. Comparing this to the mean of 83.1%, however, suggests that there are trainees with substantially lower scores to drag down the mean. In fact, on further examination, we see that only 72% of the trainees achieve the target of at least 90%. In other words, simply incorporating the random arrival pattern of procedures highlights the potential risk for some trainees to receive inadequate training opportunities, which (without intervention) could risk patient safety in the future.

3.4 Scenario 2

Finally, we add in the consideration of variation in learning speeds, assigning each resident a learning curve from a normal distribution for both procedures (means of 40 and 5 for appendectomy and mediastinoscopy, respectively). By keeping the variation in the arrival rates from the previous scenario and also increasing the standard deviation of the normal distribution of learning curves, we demonstrate that greater variability in trainee learning curves will amplify the spread of competency across a class of residents.



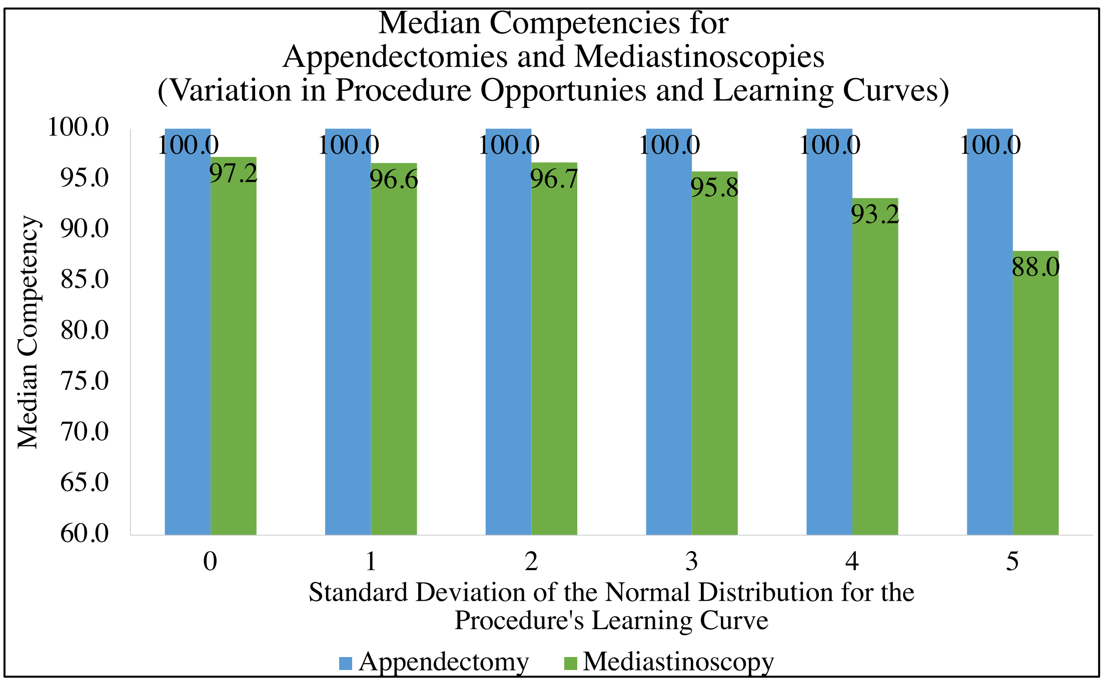


Figure 8A and 8B: Mean and median competencies for appendectomy and mediastinoscopy when there is variation in both the training opportunities and learning speeds.

With the high arrival rate of appendectomies and extensive time spent on the Acute Care service, the competency outcomes remain robust even under variation in learning curves.

However, increasing the standard deviation and thereby broadening the range of learning curves begins to significantly lower the mean competency in mediastinoscopies, even while leaving the medians relatively stable. Similar to the results of Scenario 1, it is evident that despite the majority of residents still achieving competency, greater variation in learning curves means that there is a risk of an increasing population of surgical residents who require intervention to ensure adequate preparation for independent practice. In fact, while we see that 72% achieve the target value of 90% competency with consistent learning curves, this drops to 62.7% with a standard deviation, and all the way to only 48.5% when the standard deviation is 5.

4 DISCUSSION

With the results of all three simulation scenarios, it is evident that in cases where common procedures like appendectomies have high arrival rates, the impact of variation is marginal. There are enough learning opportunities that any kind of learner can have enough practice to become competent. However, rarer procedures like mediastinoscopies are far more likely to be impacted by variable arrival rates and learning curves. While the majority of residents would still achieve the 90% competency target, greater variation leads to the growing of a population of residents that do not. Therefore, interventions are needed to ensure patient safety and prevent programs from graduating surgical residents who are fully prepared for independent practice.

A key motivator of our work is to inform ACGME guidelines, recognizing that residency programs may not adequately understand how variability could impact the competency of their trainees. While we recognize that we have made underlying assumptions about case arrivals and learning curves, it is clear that variation in competency among residents is very possible. Further research into characterizing the individual learning curves of general surgery residents and more intentionally collecting data on resident learning opportunities would allow for refined analyses that accurately represent reality and allow specific programs to adjust accordingly.

Two key areas where the simulation model could be expanded to more accurately represent surgical competency are transference and skill decay. When residents practice the basics of surgery (sutures, incisions, etc.), they are learning transferable skills that apply to many different types of surgery. For example, when a resident has the opportunity to perform a laparoscopic cholecystectomy, there is an amount of gained skill and knowledge that is applicable to performing laparoscopic colonoscopy due to the shared laparoscopic components. Residents may achieve a higher level of competency more quickly when they get to practice a surgical skill in different contexts. Skill decay refers to the fact that, due to the rigid block scheduling of residency programs, there are many months where residents do not get the opportunity to repeat a procedure that they have already learned. This factor is important because of the multiyear nature of residency programs, giving residents time during which their skills may decay (Perez, 2013).

While our research focused on assessing a potential problem within general surgery residency programs, it does not produce a clear cut solution. Rather, we highlight the impact of variability and the need to thus incorporate variability in learning programs. Key areas for future research are improved methods to measure an individual resident's surgical competency, ways to better predict outcomes given variability in both procedural opportunities and learning curves, and strategies for more dynamic block scheduling, shifting procedures to those who need more to gain competency, rather than having a one-size-fits-all approach to training. We see the importance of continued collaboration between simulation and surgical experts to identify and assess strategies for mitigating the impacts of variability in silico before piloting in clinical practice.

ACKNOWLEDGMENTS

We would like to thank the Center for Healthcare Engineer and Patient Safety (CHEPS), and the Seth Bonder Foundation for funding this research. Also, we thank Elizabeth Fisher and all CHEPS students who assisted with this work.

APPENDIX A

As seen in Figure 2, the learning curves are logistic shaped curves. The function used to calculate competency for a trainee T in procedure P in the learning curve demonstrated in Figure 2 is as follows:

$$L_{TP}(x) = 1/(1 + e^{-a(x-b)})$$

where, $L_{TP}(x)$ is the learning curve function for trainee T in procedure P, x is the number of cases completed by trainee T in procedure P, i is the LCP for trainee T in procedure P, $a = (6.7923)/i$, and $b = 0.6765 * i$.

The LCP can be used to define a function which maps the number of cases that a trainee completes in a procedure to their % competency in that procedure.

REFERENCES

Accreditation Council for Graduate Medical Education. 2022. ACGME Program Requirements for Graduate Medical Education in General Surgery. https://www.acgme.org/globalassets/pfassets/programrequirements/440_generalsurgery_2022.pdf, accessed 4th July.

Accreditation Council for Graduate Medical Education. 2019. Defined Category Minimum Numbers for General Surgery Residents and Credit Role. https://www.acgme.org/globalassets/definedcategoryminimumnumbersforgeneralsurgeryresidentsandcreditrole.pdf, accessed 10th April.

Accreditation Council for Graduate Medical Education. 2017. Case Requirements for Residents Beginning on or after July 1, 2017. https://www.acgme.org/globalassets/pfassets/programresources/caserequirementsforresidentsbeginningjuly2017.pdf, accessed 10th April.

Aziz, H., T. James, D. Ramulla, L. Sher, Y. Genyk, M.E. Sullivan, M.R. Sheikh. 2021. "Effect of COVID-19 on Surgical Training Across the United States: A National Survey of General Surgery Residents". *Journal of surgical education* 78(2):431-439.

Gart, J. J. 1975. "The Poisson Distribution: The Theory and Application of Some Conditional Tests". In *A Modern Course on Statistical Distributions in Scientific Work*, edited by G. P. Patil, S. Kotz, and J. K. Ord, 125-140. Dordrecht: Springer.

George, B. C., J.D. Bohnen JD, R.G. Williams, S.L. Meyerson, M.C. Schuller, M.J. Clark, A.H. Meier, L. Torbeck, S.P. Mandell, J.T. Mullen, D.S. Smink, R.E. Scully, J.G. Chipman, E.D. Auyang, K.P. Terhune, P.E. Wise, J.N. Choi, E.F. Foley, J.B. Dimick, M.A. Choti, N.J. Soper, K.D. Lillemoe, J.B. Zwischenberger, G.L. Dunnington, D.A. DaRosa, J.P. Fryer JP. 2017. "Readiness of US General Surgery Residents for Independent Practice". *Annals of surgery* 266(4):582-594.

Perez, R.S., A. Skinner, P. Weyhrauch, J. Niehaus, C. Lathan, S. D. Schwaitzberg, C. G. L. Cao. 2013. "Prevention of Surgical Skill Decay". *Military Medicine* 178(10): 76–86.

Subramonian, K., G. James. 2004. "The 'Learning Curve' in Surgery: What is it, How do We Measure it and Can We Influence it?". *BJU international* 93(9):1173-1174.

AUTHOR BIOGRAPHIES

FUMIYA O. ABE-NORNES is a Research Assistant at the Center for Healthcare Engineering and Patient Safety. He holds a B.S. in Biophysics from the University of Michigan. His primary research interests include using simulation in healthcare settings and mixed methods analysis of community level interventions. His email address is fumiyar@umich.edu.

SAMIR AGARWALA is a senior studying Computer Science at the University of Michigan. His primary research interests are in the field of computer vision related to scene understanding, multi-model networks and training learning-based models using limited supervision. He will be joining Stanford University as a Master's student in Fall 2022. His email address is samirag@umich.edu.

AMY E.M. COHN is the Faculty Director of the Center for Healthcare Engineering and Patient Safety (CHEPS). Her primary research interests are in applications of combinatorial optimization, particularly to healthcare and aviation, and to the challenges of optimization problems with multiple objective criteria. Her email address is amycohn@umich.edu.

RISHINDRA M. REDDY earned his undergraduate and medical degrees at Northwestern University, and completed his Thoracic Surgery residency at the University of Washington. His research interests include improving outcomes for patients with thoracic cancers and resident and medical student education. His email address is reddyrm@med.umich.edu.

NATHAN SMITH is a junior studying Movement Science in the School of Kinesiology at the University of Michigan. His primary research interests include utilizing Industrial Operations Engineering tools and techniques to tackle healthcare problems. His email address is nsmi@umich.edu.

ANGELA THELEN is a first-year research fellow at the Center for Healthcare Outcomes and Policy. She earned her medical degree at Oakland University William Beaumont School of Medicine. Her academic interests revolve around surgical education and improving surgical quality and patient outcomes through education. Her email address is aethelen@umich.edu.

RACHEL ZHANG is a junior studying Biomedical Engineering at the University of Michigan through the College of Engineering's honors program. Rachel has worked as a research assistant at the Center for Healthcare Engineering and Patient Safety since May of 2021. Her email address is rwzhang@umich.edu.

BRIAN C. GEORGE is an Assistant Professor of Surgery and serves as the Director of Educational Research at the Center for Health Outcomes and Policy as well as the Director of the Center for Surgical Training and Research at the University of Michigan. His research interest is in laying the groundwork for an ongoing national initiative in healthcare education quality improvement. His email address is bcgeorge@med.umich.edu.