

Title Page
Final Progress Report

**Determining a Learning Curve for Complex Laparoscopic
Gastrointestinal Surgery**

Principal Investigator: Dennis L. Fowler, MD, MPH
Co-Investigator: R. Todd Ogden, PhD
Co-Investigator: Y. Brenda Chang, MPH

Columbia University in the City of New York
New York, NY

Funded by the
Agency for Healthcare Quality and Research (AHRQ)

August 1, 2012-July 31, 2014

Grant Number 5R03HS021549

Structured Abstract

Purpose: The purpose of this project is to validate the Global Operative Assessment of Laparoscopic Skills (GOALS) as an assessment tool and to define the learning curve for complex laparoscopic gastrointestinal surgery.

Scope: The Fellowship Council requires that program directors assess the technical performance of Minimally Invasive Surgery fellows once each quarter during their 1-year fellowship using the GOALS tool.

Methods: The investigators collected and analyzed all of the reported GOALS scores for surgical fellows performing laparoscopic surgery during the academic year 2013-2014. We analyzed case difficulty and time of year in fellowship as potential predictors of performance. Performance scores throughout the fellowship year were graphed to create learning curves for overall performance and each of the five domains.

Results: In total, 402 performance assessments were submitted for 148 unique fellows. Overall performance, depth perception, bimanual dexterity, efficiency, tissue handling, and autonomy all improved significantly throughout the course of the fellowship year ($p < 0.001$ for each). All domains improved significantly for gastric bypass and bariatric procedures, although only overall, bimanual dexterity, and efficiency were significant for colectomies. Two predictor variables were significantly related to performance scores.

Key Words: laparoscopic surgery, technical skills, assessment, GOALS

Purpose

The purpose of this project is to validate the Global Operative Assessment of Laparoscopic Skills (GOALS) assessment tool and define the learning curve for complex laparoscopic gastrointestinal (GI) surgery. The learning curve may lead to broader implementation of proficiency-based training. The ultimate outcome of this work should be not only better-trained surgeons but also a validated method for documenting that they are in fact better trained.

The objectives of the study:

1. *To determine whether previous experience, case difficulty, previous performance scores, or time during the fellowship year impact the technical performance of fellows during complex laparoscopic GI surgery.* Using scores determined by using the Global Operative Assessment of Laparoscopic Skills (GOALS) tool, the investigators propose to test various regression models to determine which variables or combinations of variables significantly impact performance.
2. *To establish construct validity for the GOALS assessment tool for complex laparoscopic surgery.* By comparing performance scores of fellows at the beginning of the fellowship year with scores at the end of the fellowship year, and accounting for any impacting factors (previous experience, length of time as a fellow, previous performance scores, or case difficulty), we will determine whether GOALS can differentiate novice fellows from graduating fellows.
3. *To define the learning curve for complex laparoscopic surgery.* After accounting for various factors that will influence performance, we will define the learning curve by plotting the performance scores of fellows over time.

Scope

Background

Medical students are expected to demonstrate the core knowledge of surgery that is requisite of any medical practitioner. Junior surgical trainees must demonstrate their knowledge of basic medical science and clinical care for surgery patients. More advanced surgical trainees are expected to demonstrate detailed knowledge of operative surgery, surgical judgment, and competent decision making. Once in independent practice, surgeons participate in continuing professional development. Despite this, at no stage in this learning process are the technical skills of surgeons objectively assessed with a validated assessment tool.

The technical skill of the surgeon is a major factor in determining the outcome of a surgical procedure, but current certification requirements include a written examination, an oral examination, and satisfactory completion of the Fundamentals of Laparoscopic Surgery (FLS). FLS certification requires passing an assessment that measures knowledge, case/problem management skills, and basic technical skills required for laparoscopic surgery. The FLS test is widely considered to be a valid assessment of basic technical skills, but it does not measure performance in the context of an actual procedure nor does it measure advanced skills.

Traditional methods of assessment of technical skills rely upon subjective evaluations by senior staff members, case records, and on occasion, complication and mortality rates. These have been shown to be neither valid nor reliable measures of technical proficiency. Other factors – public demand, reduced resident work hours, and regulatory mandates – also have contributed to our awareness of the need to objectively assess technical proficiency in surgery at all levels of experience.

The Learning Curve

A learning curve describes the rate at which individuals acquire the skill to perform a procedure. Competence is not achieved until the learning curve flattens at a reasonably high level of performance (assessment scores stop improving). The number of cases required to reach this plateau for most learners could inform a curriculum for surgical training. A well-defined learning curve may be even more important for defining the curriculum for fellows, who must learn to perform more complex operations.

Context

The acquisition of technical skills is one of the fundamental goals of postgraduate surgical training; however, a validated objective tool to assess the technical skills of trainees remains elusive. There is a need for rigorous, reliable, and valid methods for assessing the surgical performance of residents and fellows. Only with such an assessment can progress through a curriculum be measured. Furthermore, by measuring the performance of trainees, the curriculum may be modified to ensure adequate training. Adequate surgical training has traditionally been defined by the number of years in training rather than by the achievement of proficiency as defined by objective performance measures.

The evaluation of trainees is often retrospective, leading to recall bias. However, if performed in a timely manner and performed using fixed criteria, such as a Global Rating Scale (GRS), observation by experts may be a reliable and valid evaluation tool. The Global Operative Assessment of Laparoscopic Skills (GOALS) is a tool that has been validated for measuring technical performance during less complex laparoscopic surgical procedures, such as cholecystectomy and appendectomy, but has not been validated for the evaluation of performance during complex operations. The purpose of this project is to validate GOALS as an objective measure of technical performance for complex laparoscopic gastrointestinal (GI) operations and enable its use for determining proficiency of trainees prior to completing their training. By validating this tool, we can move to proficiency-based training for laparoscopic surgery and satisfy all constituents that, by the end of their training, surgeons are proficient in these techniques. Also, with a validated tool, we have the potential to objectively assess the skills of surgeons at intervals during their careers to provide them with a reliable method for determining what deficiencies, if any, exist in their skill set.

The requisite next step is to create or identify and then validate an assessment tool. The GOALS tool exists, and the investigators have attempted to validate it for objectively assessing the performance of a surgeon during complex laparoscopic surgery.

Participants

Post-Residency Fellowship Training

Fellowship training in complex surgery is continually evolving, not only to keep up with the advances in patient care and the adoption of new technologies but also to meet the needs of trainees. The Fellowship Council oversees the training of ~180 surgical fellows who receive training during a 1-year fellowship in complex GI surgery. Valid and reliable performance assessment tools are essential to ensure that competencies are acquired. To enable program directors to provide better formative feedback regarding each fellow's technical performance, the Fellowship Council initially enabled program directors to record and voluntarily submit GOALS scores for their fellows using a web-based system. In the second year of the study, the Fellowship Council mandated reporting of technical performance at least once each quarter during the academic year. The analyses of the data from the assessments provided during the second year of this project provide the important results from this AHRQ-funded project. By analyzing the resulting data set, we started the process to identify factors that influence the fellows' learning curve and define the learning curve for complex laparoscopic GI surgery.

The Fellowship Council

The Fellowship Council was founded in 1999 as a result of the combined effort of three surgical societies: SAGES (Society of American Gastrointestinal and

Endoscopic Surgeons), SSAT (Society for Surgery of the Alimentary Tract), and AHPBA (American Hepato-Pancreato-Biliary Association). A tri-partite committee composed of representatives from each of these organizations created the council and its structure. A small number of fellowships in complex GI surgery (typically focusing on minimally invasive surgery, liver and pancreas surgery, or bariatric surgery) existed at that time, but there was no organized application process and no accreditation of these fellowships. In the interval since 1999, The Fellowship Council has developed a robust infrastructure using modern information technology for its application, accreditation, and matching processes. Currently, it oversees more than 130 fellowship programs that train nearly 180 fellows. The majority of fellowships focus on minimally invasive general surgery and/or bariatric surgery, although smaller numbers of fellowships focus on hepato-pancreatic-biliary, colorectal, or thoracic surgery, and a few fellowships focus on flexible GI endoscopy for surgeons.

The Fellowship Council has pursued efforts to both improve the educational process and better define the requirements for excellent training. Beginning in the most recent academic year, the Fellowship Council formally initiated the use of assessment tools to measure and report the technical performance of fellows. As part of this initiative, the Fellowship Council mandated reporting an assessment of at least one laparoscopic surgery case per fellow per quarter using the GOALS tool.

Methods

Study Design

In this retrospective study, data collected by the Fellowship Council for the 2013-2014 fellowship year were analyzed to validate the GOALS assessment tool. All fellows with GOALS scores in the Fellowship Council database were included in the analysis. Any fellows who did not have GOALS data in the Fellowship Council database were excluded. The GOALS scores were analyzed to determine which factors influenced the learning curve and to define the learning curve for complex laparoscopic gastrointestinal surgery. The GOALS tool is shown in Table 1. The GOALS tool includes a well-defined descriptive anchor for ratings of 1, 3, and 5 for each of the five domains.

Data Modeling

To accomplish the study goals, performance scores were analyzed using linear mixed effects with the performance score as outcome, fellow as random effect, and the difficulty of procedure and time in fellowship program as fixed effects. To allow for apparent nonlinearity of the learning curve, we log-transformed the time in the program. Previous experience for each fellow was not available and thus was not included in the model. The random effect component to the model allows us to model fellow-to-fellow variability and to account for correlation among different scores for a given subject. This also allowed us to estimate, within the estimated model, an inherent skill level for each subject.

Data Collection

The data sets used for this study contain the performance scores of fellows who are learning complex laparoscopic GI operations. The data sets were obtained from the Fellowship Council and are being used in this study with the permission of the Fellowship Council. All data points are entered into the Fellowship Council's central database by fellowship program directors. In addition to the performance scores, the reports include data on the name and date of the procedure, and the difficulty of the case as rated on a 5-point Likert scale.

The Fellowship Council mandated technical performance reporting to the central database beginning in academic year 2013-2014. This data set was obtained in the final month of this grant (July 2014). These data have been analyzed and these analyses constitute the substance of this Final Progress Report.

Table 1. Global Operative Assessment of Laparoscopic Surgery (GOALS) tool.

Domains	1	2	3	4	5
Depth perception	Constantly overshooting target, hits backstop, wide swings, slow to correct	-	Some overshooting or missing plane but corrects quickly	-	Accurately directs instruments in correct plane to target
Bimanual dexterity	Use of 1 hand, ignoring non-dominant hand, poor coordination between hands	-	Use of both hands but does not optimize interactions between hands to facilitate conduct of operation	-	Expertly utilizes both hands in a complementary manner to provide optimal working exposure
Efficiency	Uncertain, much wasted effort, many tentative motions, constantly changing focus of operation, or persisting at a task without progress	-	Slow, but planned and reasonably organized	-	Confident, efficient and safe conduct of operation, maintaining focus on component of procedure until better done by another approach
Tissue handling	Rough, tears tissue by excessive traction, injures adjacent structures, poor control of coagulation device (recoil), grasper frequently slip off	-	Handles tissues reasonably well, with some minor trauma to adjacent tissues e.g., coagulation of liver, causes unnecessary liver bleeding, occasional slipping of grasper	-	Handles tissues very well with appropriate traction on tissues and negligible injury of adjacent structures. Uses energy sources appropriately but not excessively
Autonomy	Unable to complete entire procedure, even in a straightforward case and with extensive verbal guidance	-	Able to complete operation safely with moderate prompting	-	Able to complete operation independently without prompting
Level of difficulty	Easy exploration and dissection	-	Moderate difficulty (e.g. mild inflammation, scarring, adhesions, obesity, or severity of disease)	-	Extreme difficulty (e.g., severe inflammation, scarring, adhesions, obesity, or severity of disease)

Results

The performance of each of 148 unique fellows during 402 complex laparoscopic operations was assessed using the GOALS tool for four types of operations: laparoscopic Roux-en-Y gastric bypass ($n = 87$), laparoscopic bariatric surgery ($n = 162$, including gastric bypass surgeries), laparoscopic colectomy ($n = 18$), and laparoscopic Nissen fundoplication ($n = 27$). Of the 148 fellows, 60 (41%) had only one set of performance scores available. The mean scores for each quarter for each domain increased throughout the fellowship year (Fig. 1).

Table 2. Descriptive statistics – assessment of laparoscopic cases

Procedure	Surgical records		Unique fellows	Case difficulty		
	<i>n</i>	(%)		<i>n</i>	Mean	Median
Total	402	100	148	3.67	4	1-5
Roux-en-Y gastric bypass	87	21.64	50	3.53	3	2-5
Bariatric surgery	162	40.30	78	3.46	3	1-5
Colectomy	18	4.48	13	3.78	4	2-5
Nissen fundoplication	27	6.72	22	3.56	4	1-5

*Unique fellows total more than 148, because some fellows performed more than one procedure each.

Figure 1 displays the raw scores and the estimated mean learning curve for overall performance and for performance in each of the domains using a linear mixed effects model. The domains of overall performance, depth perception, bimanual dexterity, efficiency, tissue handling, and autonomy all demonstrated statistically significant learning during the fellowship year.

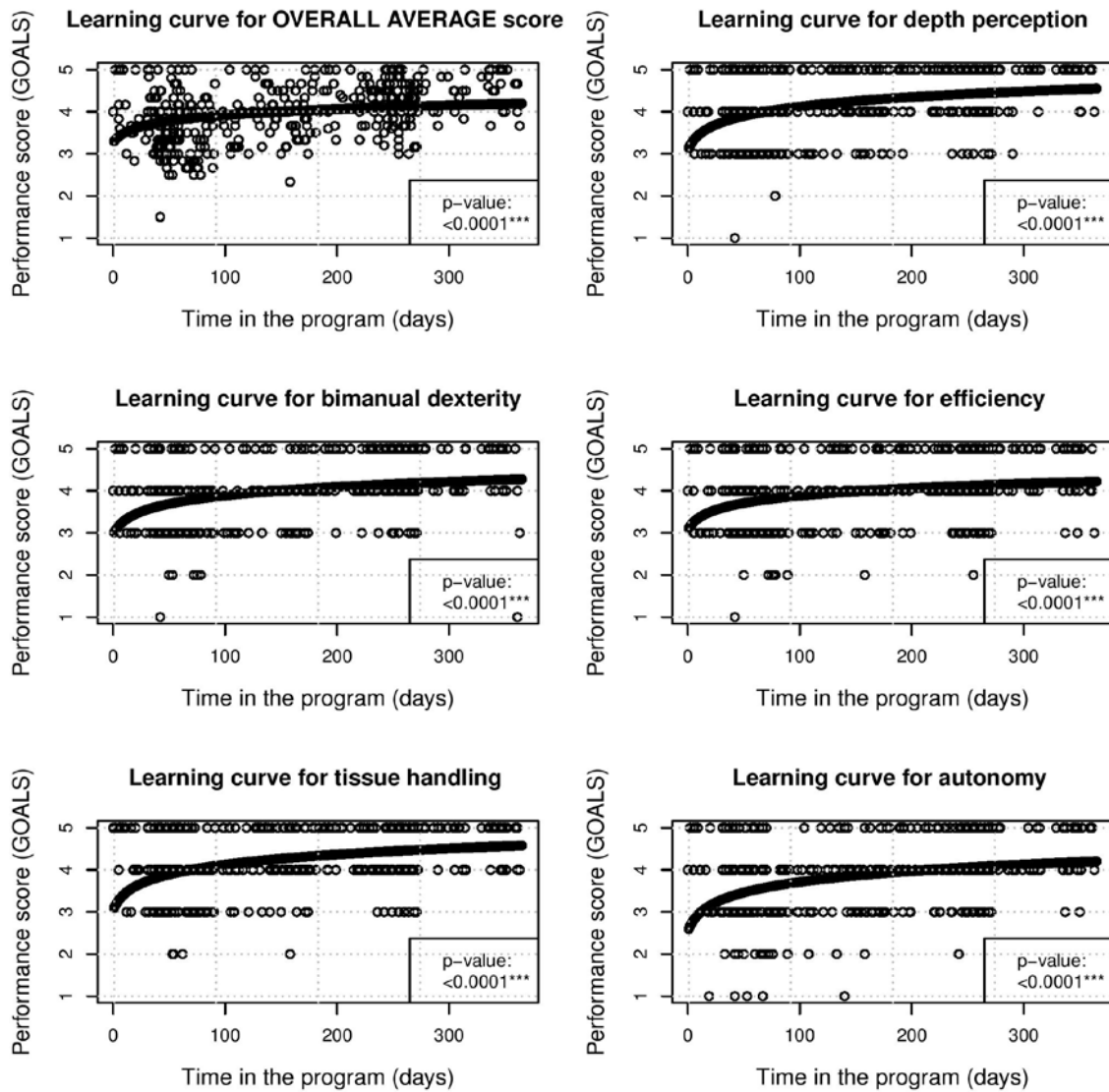


Figure 1. Learning curves for overall performance and five performance domains (GOALS)

For Roux-en-Y gastric bypass procedures and bariatric surgeries, Figures 2 and 3 illustrate, respectively, the estimated average learning curve for overall performance and for performance in each of the domains using a linear mixed-effects model. All the learning curves for Roux-en-Y gastric bypass and bariatric procedures were statistically significant.

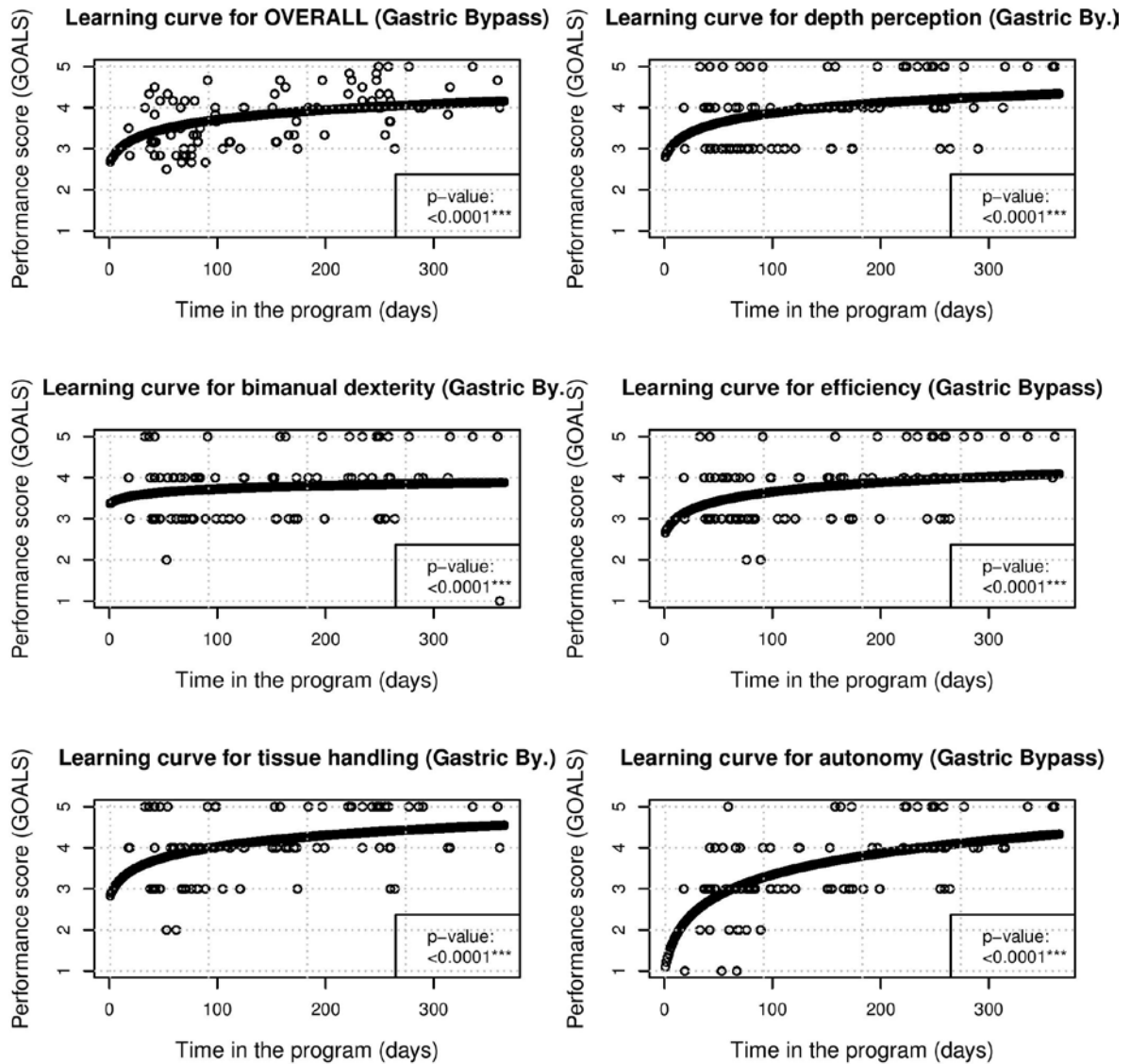


Figure 2. Learning curves for overall performance and five performance domains (GOALS) for Roux-en-Y gastric bypass procedures

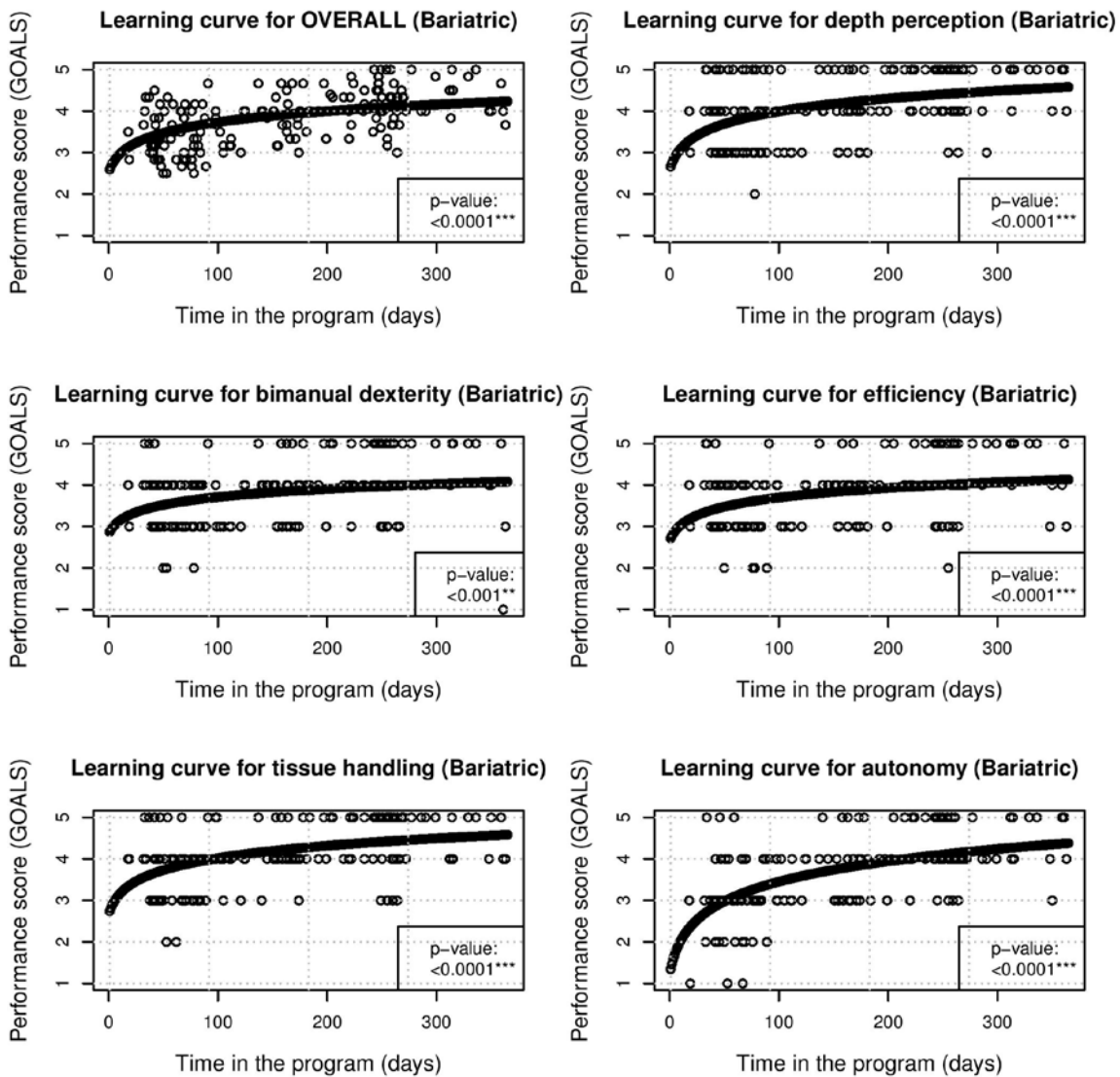


Figure 3. Learning curves for overall performance and five performance domains (GOALS) for bariatric procedures

For colectomy procedures, the learning curves for overall performance, bimanual dexterity, and efficiency demonstrated statistically significant learning during the fellowship year. Although the learning curves for depth perception and tissue handling showed a trend toward improvement as illustrated in Figure 4, the curves were not statistically significant. Not only was the number of assessments for the year quite small ($n = 18$), there was only one assessment reported in the 4th quarter of the academic year.

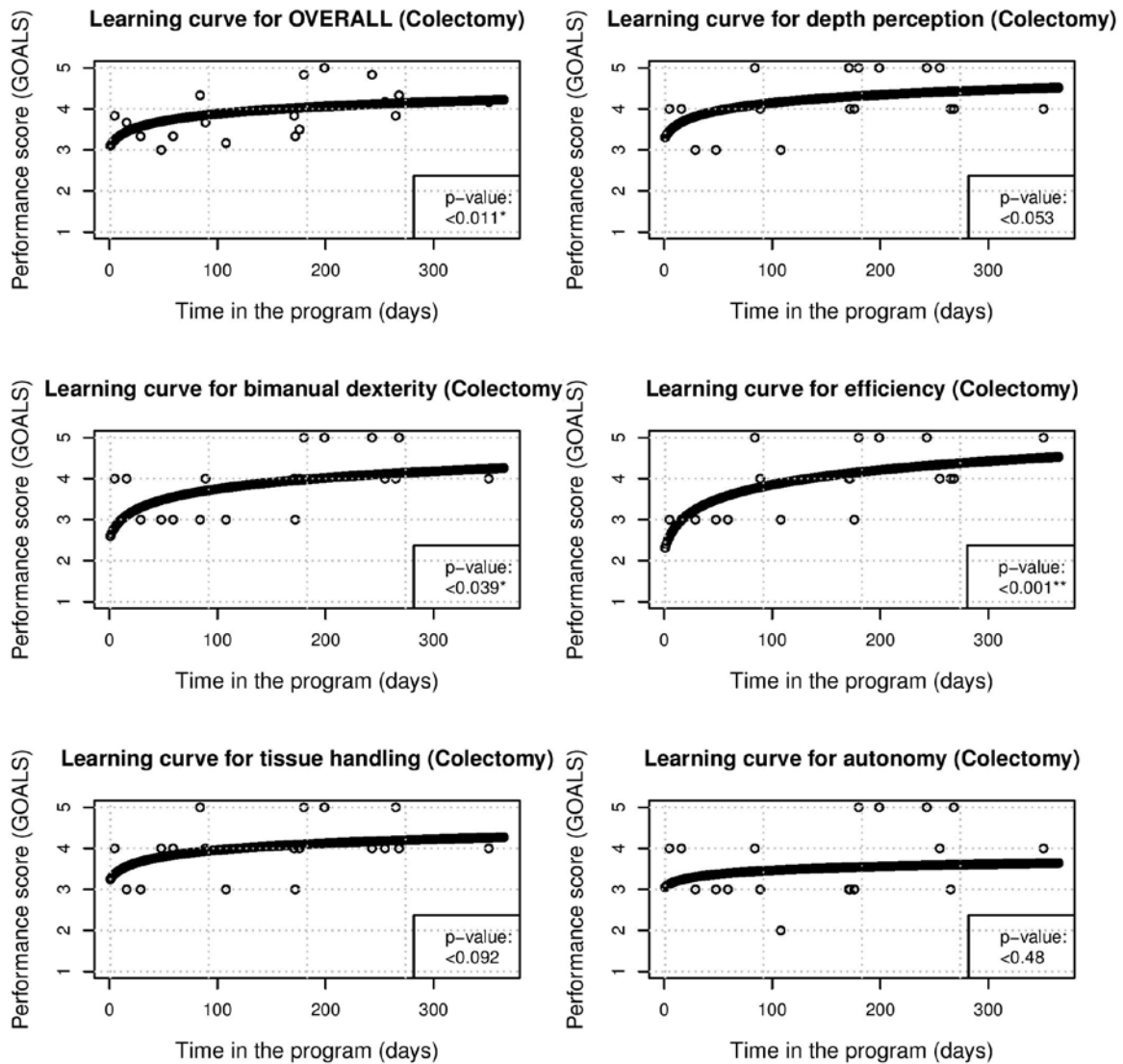


Figure 4. Learning curves for overall performance and five performance domains (GOALS) for colectomy procedures

For Nissen fundoplication procedures, none of the learning demonstrated statistically significant learning during the fellowship year. Although the learning curves for overall average performance, depth perception, and tissue handling showed a trend toward improvement as illustrated in Figure 5, the curves were not statistically significant. The number of assessments for the year was small ($n = 27$), and the number of assessments per unique fellow was very small, suggesting that the assessment scores later in the year were not for the same fellows as the scores in the early part of the year.

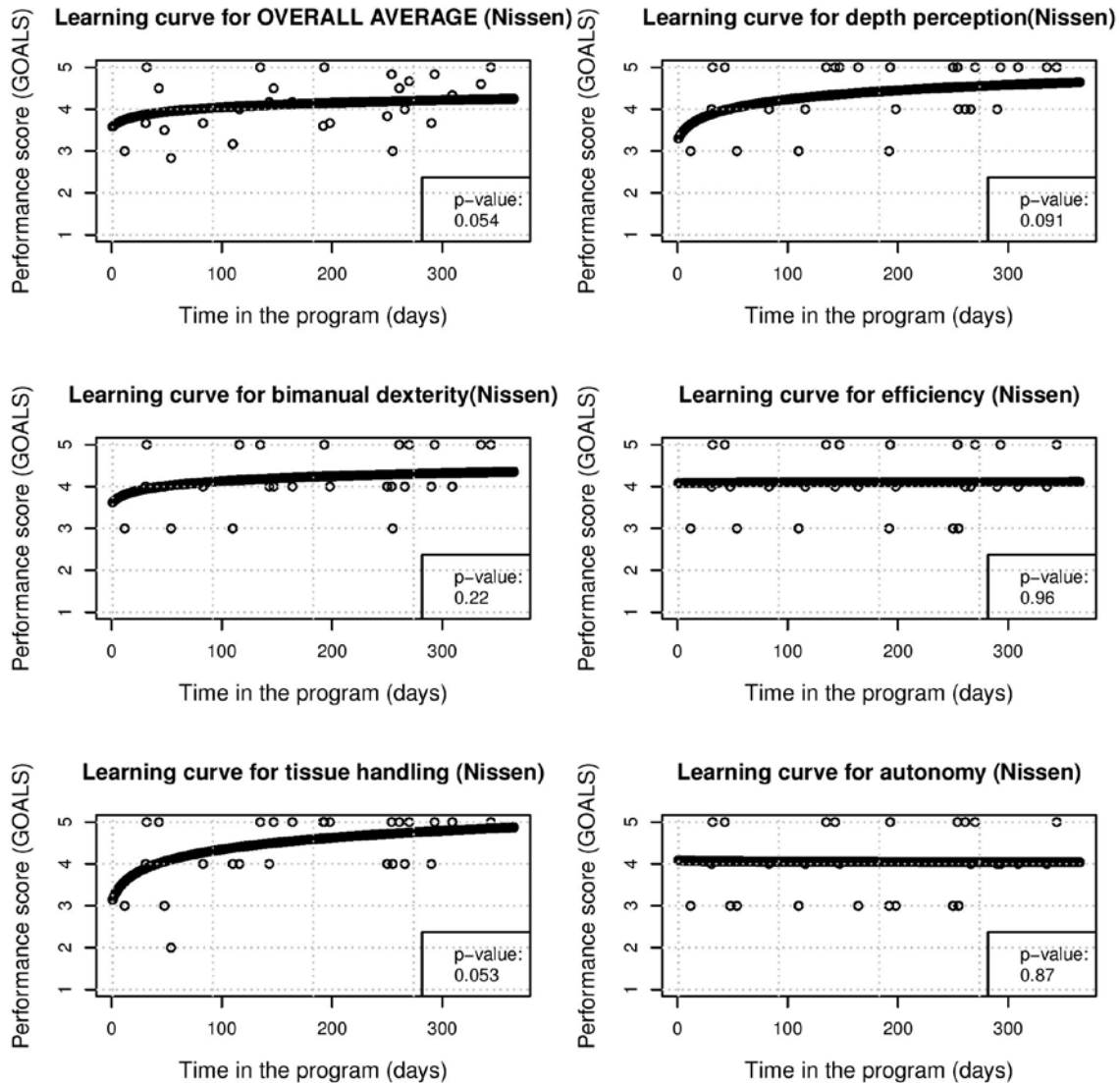


Figure 5. Learning curves for overall performance and five performance domains (GOALS) for Nissen fundoplication procedures

In summary, the GOALS assessment tool documents statistically significant improvement in the performance of fellows during complex laparoscopic surgery across the time of the fellowship year of training. The assessment tool demonstrates construct validity because it documents significantly better performance in the fourth quarter than in the first quarter, although the improvement in performance is gradual throughout the course of the fellowship year. The fellows entered the year with a moderate degree of skill, consistent with what would be expected after the fellows have completed a 5-year surgery residency. The learning curves demonstrate progression from this moderate degree of technical skill at the beginning of the year to a superior degree of technical skill at the end of the year. Future work can more precisely define the learning curve for each specific operation if there are a larger number of assessments.

List of Publications and Products

Hogle NJ, Liu Y, Ogden RT, Fowler DL. Evaluation of surgical fellows' laparoscopic performance using global operative assessment of laparoscopic skills (GOALS). Surg Endosc. 2014 Apr;28(4):1284-90. PMID: 24414454