Preliminary Evaluation of a Laparoscopic Common Bile Duct Simulator for Pediatric Surgical Education

Ben Schwab, MD,[1] Deborah M. Rooney, PhD,[2] Eric S. Hungness, MD,[1] and Katherine A. Barsness, MD, MS[3]

Abstract

Purpose: Laparoscopic common bile duct exploration (LCBDE) decreases overall costs and length of stay in patients with choledocolithiasis. However, utilization of LCBDE remains low. We sought to evaluate a previously developed general surgery LCBDE simulator among a cohort of pediatric surgical trainees. The study purpose was to evaluate the content validity of an LCBDE simulator to support or refute its use in pediatric surgery education.
Materials and Methods: After IRB exempt determination, 30 participants performed a transcystic LCBDE using a previously developed simulator and evaluated the simulator using a self-reported 28-item instrument. The instrument consisted of two primary domains (Quality and Ability to Perform) that were rated using twenty-five 4-point rating scales and one 4-point global rating scale. Validity evidence relevant to test content was evaluated using a many-facet Rasch model. Interitem consistency was estimated using Cronbach's alpha. $P < .05$ was considered statistically significant.
Results: The highest combined observed averages were for the Value subdomain (OA = 3.79), whereas the lowest ratings were for the Physical/visual attributes subdomain (OA = 3.19). The averaged global rating was 3.14, consistent with this simulator can be considered for use in pediatric LCBDE training, but could be improved slightly. Rasch indices were favorable and supported evidence relevant to test content. Interitem consistency estimates were also favorable, with α values of 0.94 and 0.56 for Qualities and Ability, respectively.
Conclusions: Overall, participants rated the LCBDE simulator highly valuable for pediatric surgical education and felt that it could be used as an educational tool with minor modifications.

Keywords: choledocholithiasis, simulation, education, pediatric, surgery

Introduction

THE FINDING OF CHOLEDOCHOLITHIASIS in patients who present with cholelithiasis presents a management dilemma for the treating surgeon, namely what is the optimal method for achieving clearance of ductal obstruction? The traditional approach consists of endoscopic retrograde cholangiopancreatography (ERCP) as a means of alleviating the ductal obstruction before surgical removal of the gallbladder. However, laparoscopic common bile duct exploration (LCBDE) has been proven to provide excellent clearance rates and periprocedural morbidity equivalent to that of ERCP, with the added benefit of shorter hospital stays and reduced cost.[1–4] Similar studies have demonstrated the efficacy and safety of the combined approach in the management of choledocholithiasis in children.[5,6]

However, LCBDE is a procedure that requires advanced laparoscopic skills, which may explain the low utilization of the procedure.[7] To address this educational gap, an LCBDE simulator was developed, evaluated, and incorporated into a mastery learning curriculum used for training senior surgical residents at our institution.[8] Universal attainment of the mastery standard among residents was achieved with a comprehensive curriculum and deliberate practice on the simulator.[9]

In this study, we sought to evaluate the clinical utility of the previously developed LCBDE simulator for use in training of pediatric surgery trainees, with an emphasis on assessment of the simulator's content validity, and to determine whether or not physical modifications to the pre-existing simulator would be necessary before it could be used in pediatric surgical education.

Materials and Methods

Study participants

Following review and exempt determination from Northwestern University's Institutional Review Board, 31 second-

[1]Department of Surgery and Medical Education, Northwestern University Feinberg School of Medicine, Chicago, Illinois.
[2]Department of Learning Health Sciences, University of Michigan Medical School, Ann Arbor, Michigan.
[3]Department of Surgery and Medical Education, Ann and Robert H. Lurie Children's Hospital of Chicago, Northwestern University Feinberg School of Medicine, Chicago, Illinois.

year pediatric surgery trainees participated in a 2-day pediatric surgery course hosted by Northwestern Simulation (Chicago, IL) in September of 2015. During the course, 30 participants (96.8%), representing more than half of all pediatric surgery programs in the United States and Canada, evaluated the LCBDE simulator.

LCBDE simulator

The previously developed LCBDE simulator is a self-contained model of the liver, gallbladder, extrahepatic biliary system, and duodenum (created out of purely synthetic materials) that is then placed inside a standard Fundamentals of Laparoscopic Surgery (FLS) box trainer (VTI Medical, Waltham, MA).[8] This arrangement allows for the use of the FLS camera to simulate the laparoscopic view obtained from a periumbilical camera port. A second video camera system provides a simulated real-time fluoroscopic view that is displayed on a second monitor and is controlled by the user with a foot pedal. A fiber-optic or video choledocoscope (Karl-Storz, Tuttlingen, Germany) provides an endoscopic view that is displayed in a picture-in-picture manner in conjunction with the laparoscopic view on a single monitor (Fig. 1).

The laparoscopic instruments and endoscopic equipment (Cook Medical, Bloomington, IN) necessary to perform a complete transcystic LCBDE were provided. Participants were then asked to retrieve a 6-mm multifaceted bead, simulating an impacted gallstone in the common bile duct.

Assessment of simulator

After completing the simulated LCBDE, participants evaluated the simulator using a paper instrument. The 26-item instrument consisted of two targeted domains; simulator qualities (Qualities), and ability to perform LCBDE-relevant tasks (Ability). Additionally, there were two demographic items targeting participants' familiarity with LCBDE-relevant

FIG. 1. LCBDE simulator views **(A)** Laparoscopic, **(B)** Endoscopic, and **(C)** Fluoroscopic. LCBDE, laparoscopic common bile duct exploration.

equipment and setup and current comfort level with performing an LCBDE for choledocolithiasis (both scored 1 = Not comfortable to 3 = Very comfortable).

Participants evaluated the characteristics and qualities of the simulator (Qualities) across three subdomains (Physical/visual attributes, Realism of experience, and Value). The four items targeting Physical/visual attributes and 11 items targeting Realism of experience were rated using 4-point rating scales ranging from 1 (Not realistic) to 4 (Highly realistic), while the four Value items, including relevance to practice and value of the simulator as a training and a testing tool, were rated on 4-point scales (with added do not know option) and a global rating (4-point rating scale). Participants also rated their personal ability to perform seven individual tasks during the LCBDE procedure (Ability) using a 4-point rating scale, ranging from 1 (Very difficult to perform) to 4 (Very easy to perform).

Statistical analysis

Validity evidence relevant to test content and internal structure was evaluated using indices from a modern measurement model, while additional evidence of internal structure (interitem consistency) was estimated using Cronbach's alpha. These types of evidences are described below.

Evidence relevant to test content.   To evaluate validity evidence relevant to test content, we employed an application from modern test theory: a Rasch model.[10] Analysis was performed using the Facets software v. 3.68.2 (Linacre, 2011). For this study, we applied a many-facet Rasch model consisting of four facets (participants × comfort with equipment × self-efficacy × items) to acquire three indices used to evaluate content validity—observed averages (OAs), item outfit statistics, and point–measure correlations. These indices, described in greater detail by previous work,[11] were adapted from Wolfe and Smith[12] and are summarized as follows:

Observed averages.   The OA for each of the 15 items indicates the participants' averaged ratings. Higher OAs suggest that the perceived representativeness and the perceived realism of the simulator's features are high, while lower OAs suggest lower representativeness.

Item outfit statistics.   As described by Linacre, item mean square outfit (Outfit MS) statistics show the size of the randomness or variability in items' ratings.[13] With expected value of 1.0, values <1.0 suggest that ratings are predictable (in high agreement), while values greater than 1.0 indicate unpredictability (highly variable). In this study, we considered the existence of items with Outfit MS values higher than 2.0 a threat to content validity.

Point–measure correlations.   The point–measure correlation, also called the item–measure correlation, provides a correlation that identifies the degree in which the scores on an item are consistent with the average scores of the remaining items. Positive point–measure correlations are ideal and indicate that items contribute useful information to the construct measured by the test as a whole. For this application, a negative value for a particular item may suggest that the item may be measuring a different construct than the other items and fails to offer evidence of content validity.
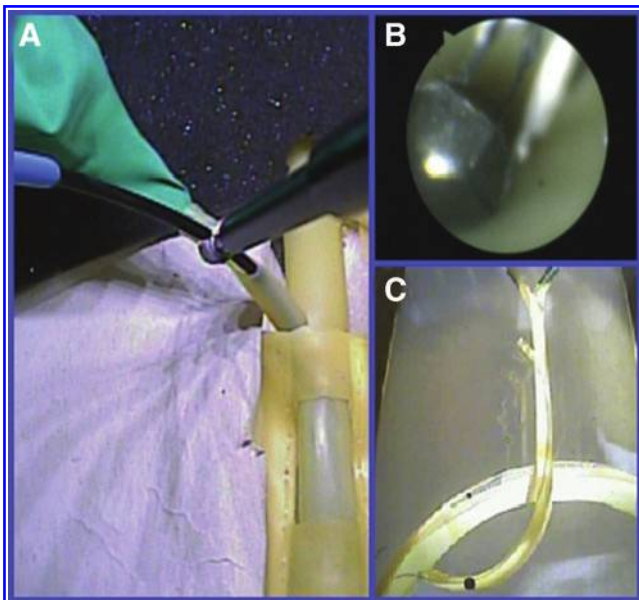
Evidence relevant to internal structure. Using a traditional method based on classical test theory, we evaluated interitem consistency estimated by Cronbach's alpha for the two primary domains—simulator quality and participants' ability to perform tasks. With a possible range of 0.0–1.0, an acceptable internal consistency estimate (0.70) would suggest that the combined items adequately measure the single intended construct. This, considered in combination with nonextreme ($|\pm 2.0|$) item Outfit MS statistics, would support evidence of internal structure.

Results

The majority of participants self-reported a high familiarity with, and comfort toward, the use of LCBDE in the pediatric population. Twenty participants (66.7%) self-reported their familiarity with LCBDE-relevant equipment as at least somewhat comfortable, while 22 participants (77.3%) self-reported their ability to successfully complete an LCBDE on a pediatric patient as at least somewhat comfortable. There were no statistical differences at item or domain levels when comparing ratings of participants with low familiarity and/or comfort with those with high familiarity/comfort ($P = .13, .92$). Given these data, findings are reported as combined OAs.

Evidence relevant to test content

Findings indicated that the highest combined OAs were for the Qualities–Value subdomain (OA = 3.79), whereas the lowest ratings were for the Qualities–Physical/visual attributes subdomain (OA = 3.19) (Table 1). Lowest rated items were Performing intraoperative cholangiogram and Realism of balloon dilation (OA = 2.96 and 3.00, respectively). The

averaged global rating was 3.14, consistent with this simulator can be considered for use in pediatric LCBDE training, but could be improved slightly. Participants' self-reported ability to complete the seven LCBDE tasks was high, indicated by high OAs [3.25, 3.38], aligning with somewhat easy to perform (Table 2).

For both Qualities and Ability domains, all outfit MS values fell below the acceptable threshold of 2.0, [0.41, 1.84]. The lowest outfit MS value was associated with Qualities—Expected overall experience of LCBDE in 15-year-old child, indicating a high degree of agreement with participants' OA of 3.30, which aligned with Adequate realism, but could be improved. The highest outfit MS value was associated with Qualities—Relevance of simulator to practice, indicating a high degree of variability in participants' perceived relevance of the simulator to their own practice, in spite of the high observed average (OA = 3.84), which aligned with Has a great deal of relevance. A review of the Ability domain's outfit MS indices indicated that six of the seven items had values over 1.36, suggesting relatively high variability in participants' self-reported ability to perform each of the required tasks. In spite of this finding, all indices were well under the threshold of 2.0, indicating reasonable variability.

Evidence relevant to internal structure

Point–measure correlations: Analysis of all 26 items of Qualities and Ability domains indicated that all items had positive point–measure correlations (Tables 1 and 2). For the 19 items of the Qualities domain, items ranged from 0.19 to 0.76. The point–measure correlations for the seven items of the Ability domain were lower, ranging from 0.17 to 0.48. Positive point–measure correlations for the 26 items suggest

TABLE 1. SUMMARY OF VALIDITY EVIDENCE ASSOCIATED WITH SIMULATOR QUALITIES (QUALITIES)

Item No.	Subdomain/items	Observed average (n = 30)	Outfit MS	Point–measure correlations
Physical/visual attributes		3.19	—	—
1	Laparoscopic appearance of the abdominal cavity (scale)	3.20	1.28	0.48
2	Fluoroscopic view of the biliary system	3.13	0.96	0.60
3	Endoscopic appearance of the biliary system (color, scale)	3.28	1.18	0.50
4	Endoscopic appearance/feel of common bile duct stone	3.14	0.78	0.64
Realism of experience		3.52	—	—
5	Amount of resistance needed to gain wire access to cystic duct	3.21	0.61	0.69
6	Obtaining/maintaining guidewire access	3.32	0.62	0.66
7	Performing balloon dilation of cystic duct	3.19	0.52	0.73
8	Choledocoscope	3.30	0.81	0.64
9	Basket retrieval of common duct stone	3.32	0.56	0.69
10	Performing cholangiogram	2.96	0.77	0.61
11	Realism of balloon dilation	3.00	0.94	0.63
12	Placement/manipulation of choledocoscope	3.19	0.74	0.72
13	Realism of choledocoscopic view	3.33	0.59	0.68
14	Realism of retrieval of common bile duct stone using nitinol basket	3.31	0.47	0.73
15	Expected overall experience of LCBDE in 15-year-old child	3.30	0.41	0.76
Value		3.79	—	—
16	Value of simulator as training tool	3.83	0.89	0.25
17	Value of simulator as testing tool	3.71	0.93	0.32
18	Relevance of simulator to practice	3.83	1.84	0.19
19	Global rating	3.14	1.77	0.48

LCBDE, laparoscopic common bile duct exploration; MS, mean square.

Table 2. Summary of Validity Evidence Associated with Participants' Ability
to Perform LCBDE-Relevant Tasks (Ability)

Item No.	Items	Observed average (n = 30)	Outfit MS	Point–measure correlations
	Ability to perform task	3.34		
1	Perform intraoperative cholangiogram	3.38	1.54	0.17
2	Identify CBD anatomy/location of biliary obstruction fluoroscopically	3.36	1.45	0.35
3	Obtain/maintain guidewire access to CBD	3.41	1.40	0.48
4	Perform balloon dilation of cystic duct	3.30	1.36	0.48
5	Insert choledocoscope into cystic duct and navigate into CBD	3.29	1.44	0.37
6	Identify CBD obstruction endoscopically	3.36	1.04	0.48
7	Perform basket retrieval of CBD stone	3.25	1.67	0.36

CBD, common bile duct; LCBDE, laparoscopic common bile duct exploration; MS, mean square.

that these items contribute to a single construct and offer evidence of content validity. Interitem consistency for the 19 items used to measure simulator quality was high ($\alpha = 0.94$), while interitem consistency for the seven items used to measure participants' ability to perform tasks was lower ($\alpha = 0.56$), but adequate for this preliminary study.

Discussion

In this study, we sought to evaluate a previously developed adult LCBDE simulator in regard to its applicability for training of pediatric surgery fellows to perform the task in adolescent patients. After performing an LCBDE procedure on the simulator, participants rated it as highly realistic and relevant to the needs of pediatric surgery trainees.

Childhood obesity continues to be a major health concern in the United States; current estimates of the prevalence of overweight (body–mass index [BMI] >85%) and obesity (BMI >95%) in children are 17% and 31%, respectively.[14] While a significant proportion of gallstone disease in the pediatric population is attributed to children with hemolytic disorders, multiple studies have identified obesity as a major risk factor in the development of symptomatic biliary lithiasis, leading to an increase in the number of cholecystectomies being performed in children.[15,16]

The incidence of cholelithiasis in the adult population has also been steadily increasing, mirroring the troubling upward trend in adult obesity rates. As a result of these trends, an estimated 403,000 cholecystectomies are performed in the United States annually.[17] Contemporary data suggest that 5%–17% of patients who undergo a laparoscopic cholecystectomy (LC) will be found to have choledocolithiasis intraoperatively.[18,19] Numerous randomized prospective trials have confirmed that the use of a single-stage procedure (LC+LCBDE) results in equivalent common bile duct (CBD) stone clearance rates with the added benefit of a shorter hospital stay and improved cost-effectiveness when compared with the two-stage approach (LC+ERCP).[1–4] Similar studies have confirmed these findings in pediatric populations.[5,6]

Despite these data, the clinical utilization of ERCP far exceeds that of LCBDE or open CBDE, with recent estimates suggesting that ERCP is chosen 93% of the time compared with 7% of cases managed surgically.[7] One explanation for

this large disparity is increasing surgeon unfamiliarity with operative management of the biliary tract.[20]

To address the gap in surgical training, a low-cost LCBDE simulator incorporating the laparoscopic, endoscopic, and fluoroscopic views was developed and evaluated.[8] Utilizing the simulator, a curriculum based on achievement of the mastery standard was developed and tested on the general surgery residents at Northwestern University that demonstrated universal achievement of the mastery standard among senior surgical trainees after implementation of a didactic curriculum and deliberate practice on the LCBDE simulator.[9] Use of the mastery learning standard in simulation-based education has been shown to have implications for clinical practice, including improved patient outcomes.[21]

For the purpose of this study, we focused on the evaluation of validity evidence relevant to test content and internal structure. The relatively high Outfit MS values associated with the Ability domain suggest that participants' ability was variable, but not extreme. Taking the OAs into account (all aligning with somewhat easy to perform), variability in self-reported ability levels seems to reflect an authentic skill variability among participants that parallels studies performed in the past with participants with varying degrees of experience.[22–24]

There are a number of limitations in the interpretation and application of the findings in this study. The first limitation is associated with the small sample size. Although a sample of 30 is considered adequate for low-stakes settings such as this,[25] a larger sample size would have increased the precision of the measures. The second limitation is associated with homogeneity and composition of the sample. Although the participants were from a number of institutions, they consisted primarily of trainees. With the exception of one participant who had self-reported performing 10 LCBDE pediatric cases, the majority of participant raters (27, 90%) had not performed more than two pediatric LCBDE cases before evaluating the simulator. The narrow range of experience may have decreased the variability of ratings or may not reflect authentic ratings from expert raters.

In spite of these limitations, we highlight that there were no statistically significant differences when comparing the Qualities ratings from participants with high and low self-efficacy, suggesting that ratings of the simulator's qualities were not dependent on participants' comfort with the equipment, nor their self-efficacy toward performing LCBDE on pediatric patients.

We speculate these findings would be consistent if the study was expanded to include a broad sample of experienced pediatric surgeons.

In conclusion, participants rated the LCBDE simulator highly valuable for pediatric surgical education. Initial validity evidence relevant to test content and internal structure suggests that the LCBDE simulator could be used as an educational tool with minor alterations. Based on these findings, pediatric modifications to the existing simulator are ongoing.

Disclosure Statement

No competing financial interests exist.

References

1. Bansal VK, Misra MC, Rajan K, Kilambi R, Kumar S, Krishna A, Kumar A, Pandav CS, Subramaniam R, Arora MK, Garg PK. Single-stage laparoscopic common bile duct exploration and cholecystectomy versus two-stage endoscopic stone extraction followed by laparoscopic cholecystectomy for patients with concomitant gallbladder stones and common bile duct stones: A randomized controlled trial. Surg Endosc 2014;28:875–885.
2. Rogers SJ, Cello JP, Horn JK, Siperstein AE, Schecter WP, Campbell AR, Mackersie RC, Rodas A, Kreuwel HT, Harris HW. Prospective randomized trial of LC+LCBDE vs ERCP/S+LC for common bile duct stone disease. Arch Surg 2010;145:28–33.
3. Rhodes M, Sussman L, Cohen L, Lewis MP. Randomised trial of laparoscopic exploration of common duct versus postoperative endoscopic retrograde cholangiography for common bile duct stones. Lancet 1998;351:159–161.
4. Noble H, Tranter S, Chesworth T, Norton S, Thompson M. A randomized, clinical trial to compare endoscopic sphincterotomy and subsequent laparoscopic cholecystectomy with primary laparoscopic bile duct exploration during cholecystectomy in higher risk patients with choledocholithiasis. J Laparoendosc Adv Surg Tech 2009;19:713–720.
5. Scott SS, Philip KF, Nguyen N, Liu Q, Berel D, Wang KS. Laparoscopic common bile duct exploration in children is associated with decreased cost and length of stay: Results of a two-center analysis. J Pediatr Surg 2013;48:215–220.
6. Hill SJ, Wulkan ML, Parker PM, Jones TK, Heiss KF, Clifton MS. Management of the pediatric patient with choledocholithiasis in an era of advanced minimally invasive techniques. J Laparoendosc Adv Surg Tech 2014;24:38–42.
7. Poulose BK, Arbogast PG, Holzman MD. National analysis of in-hospital resource utilization in choledocholithiasis management using propensity scores. Surg Endosc 2006;20:186–190.
8. Santos BF, Reif TJ, Soper NJ, Nagle AP, Rooney DM, Hungness ES. Development and evaluation of a laparoscopic common bile duct exploration simulator and procedural rating scale. Surg Endosc 2012;26:2403–2415.
9. Teitelbaum EN, Soper NJ, Santos BF, Rooney DM, Patel P, Nagle AP, Hungness ES. A simulator-based resident curriculum for laparoscopic common bile duct exploration. Surgery 2014;156:880–893.
10. Rasch, G. Probabilistic models for some intelligence and attainment tests, ed. C.D.I.f.E. Research (1960/1980) Chicago, IL: The University of Chicago Press.
11. Barsness KA, Rooney DM, Davis LM. Collaboration in simulation: The development and evaluation of a novel thoracoscopic neonatal simulator. J Pediatr Surg 2013;48:1232–1238.
12. Wolfe EW, Smith EV, Jr. Instrument development tools and activities for measure validation using Rasch models: Part II—Validation activities. J Appl Meas 2007;8:204–234.
13. Linacre JM. Facets Software v. 3.68. Chicago, IL: MESA Press, 2011.
14. Ogden CL, Carroll, MD, Curtin LR, Lamb MM, Flegal KM. Prevalence of high body mass index in US children and adolescents, 2007–2008. JAMA 2010;303:242–249.
15. Mehta S, Lopez ME, Chumpitazi BP, Mazziotti MV, Brandt ML, Fishman DS. Clinical characteristics and risk factors for symptomatic pediatric gallbladder disease. Pediatrics 2012;129:e82–e88.
16. Walker SK, Maki AC, Cannon RM, Foley DS, Wilson KM, Galganski LA, Wiesenauer CA, Bond SJ. Etiology and incidence of pediatric gallbladder disease. Surgery 2013;154:927–933.
17. Hall MJ, DeFrances CJ, Williams SN, et al. National Hospital Discharge Survey: 2007 Summary. National Health Statistics Reports; No. 29. Hyattsville, MD: National Center for Health Statistics, 2010.
18. Houdart R, Perniceni T, Darne B, Salmeron M, Simon JF. Predicting common bile duct lithiasis: Determination and prospective validation of a model predicting low risk. Am J Surg 1995;170:38–43.
19. Collins C, Maguire D, Ireland A, Fitzgerald E, O'Sullivan GC. A prospective study of common bile duct calculi in patients undergoing laparoscopic cholecystectomy: Natural history of choledocholithiasis revisited. Ann Surg 2004;239:28–33.
20. Helling TS, Khandelwal A. The challenges of resident training in complex hepatic, pancreatic, and biliary procedures. J Gastrointest Surg 2008;12:153–158.
21. Mcgaghie WC, Issenberg SB, Barsuk JH, Wayne DB. A critical review of simulation-based mastery learning with translational outcomes. Med Educ 2014;48:375–385.
22. Barsness KA, Rooney DM, Davis LM. The development and evaluation of a novel thoracoscopic diaphragmatic hernia repair simulator. J Laparoendosc Adv Surg Tech 2013;23:714–718.
23. Barsness KA, Rooney DM, Davis LM, O'Brien E. Preliminary evaluation of a novel thoracoscopic infant lobectomy simulator. J Laparoendosc Adv Surg Tech 2015;25:429–434.
24. Barsness KA, Rooney DM, Davis LM, O'Brien E. Evaluation of three sources of validity evidence for a synthetic thoracoscopic esophageal atresia/tracheoesophageal fistula repair simulator. J Laparoendosc Adv Surg Tech 2015;25:599–604.
25. Linacre JM. Sample size and item calibration stability. Rasch Measurement Transactions 1994;7:328.

Address correspondence to:
Katherine A. Barsness, MD, MS
Department of Surgery and Medical Education
Ann and Robert H. Lurie Children's Hospital of Chicago
Northwestern University Feinberg School of Medicine
225 E. Chicago Avenue, Box 63
Chicago, IL 60611

E-mail: kbarsness@luriechildrens.org