Systematic Review / Meta-analysis

ANNALS OF MEDICINE & SURGERY

OPEN

The use of simulators to acquire ERCP skills: a systematic review

Konstantinos Georgiou, MD, PhD[a], Kiril T. Atliev, MD[b], Ninos Oussi, MD, PhD[e,g], Nikola Boyanov, MD[c], Gabriel Sandblom, MD, PhD[f], Lars Enochsson, MD, PhD[d,h,*]

Background and Aims: Endoscopic retrograde cholangiopancreatography (ERCP) is a technically demanding diagnostic and therapeutic endoscopic procedure with a high risk for adverse events such as post-ERCP pancreatitis and bleeding. Since endoscopists with less experience have higher adverse event rates, the training of new residents on ERCP simulators has been suggested to improve the resident's technical skills necessary for ERCP. However, there is a lack of consensus on whether the training program should focus on a threshold number of procedures or be more tailored to the individual's performance. Furthermore, there is also disagreement on which form of simulator(s) should be used. Therefore, the primary outcome of this systematic review was to study the extent to which simulators used for ERCP training are correctly validated.

Methods: In 2022, a systematic search of the literature was conducted on MEDLINE and SCOPUS under the guidance of the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) 2020 protocol seeking articles with the MeSH terms 'Endoscopic Retrograde Cholangiopancreatography' OR 'ERCP' in combination with 'simulation' OR 'simulator'.

Results: The search resulted in 41 references. A total of 19 articles met the inclusion criteria and were included in the qualitative analysis. Only one of the articles fulfilled the criteria of a robust validation study.

Conclusions: Since only one of the 19 articles met the requirements for a thorough and correct validation, further studies with sufficient numbers of subjects, that evaluate complete preclinical training programs based on validated ERCP simulators are warranted.

Keywords: ERCP, simulation, systematic review, validation

Introduction

Despite developing innovative diagnostic and therapeutic endoscopic techniques that require advanced technical skills, the lack of a standardized education curriculum remains. Furthermore, most studies on endoscopic simulator training have concentrated on gastroscopy and colonoscopy[1]. Therefore, there is a need for training of more advanced endoscopic procedures such as endoscopic retrograde cholangiopancreatography (ERCP) and endoscopic ultrasound [2].

ERCP is a technically demanding procedure with a high risk for serious adverse events of which post-ERCP pancreatitis (PEP) is the most frequent (3.5–5%)[3,4]. The risk of developing PEP can be assessed relative to several patient-related variables including sex, age, and previous history of PEP, and technical factors such as injection of contrast into the pancreatic duct and difficult cannulation[5]. Furthermore, due to the development of modern diagnostic modalities such as MRCP, the use of ERCP as a diagnostic tool has decreased, which is why ERCP is now mainly a therapeutic procedure to remove common bile duct stones, take biopsies of suspected malignant stenoses, or to relieve obstructive

HIGHLIGHTS

- Endoscopic retrograde cholangiopancreatography (ERCP) is a technically demanding diagnostic and therapeutic procedure with a high risk of adverse events.
- Endoscopists with less experience present higher adverse event rates.
- Training in validated ERCP simulators could be a way to improve training and patient safety.
- This systematic review shows a lack of properly validated ERCP simulators.
- Further studies with a sufficient number of subjects and with proper validation of ERCP simulators are warranted.

[a]First Department of Propaedeutic Surgery, Hippokration General Hospital of Athens, Athens Medical School, National and Kapodistrian University of Athens, Athens, Greece, [b]Department of Urology and General Medicine, Faculty of Medicine, Medical University of Plovdiv, [c]Medical Simulation Training Center at Research Institute of Medical University of Plovdiv, Plovdiv, Bulgaria, [d]Department of Surgical and Perioperative Sciences, Surgery, Sunderby Research Unit, Umeå University, Umeå, [e]Centre for Clinical Research, Region Sörmland, Uppsala University, Eskilstuna, [f]Department of Clinical Science and Education Södersjukhuset, [g]Division of Urology and [h]Division of Surgery, CLINTEC, Karolinska Institute, Stockholm, Sweden

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

*Corresponding author. Address: Department of Surgical and Perioperative Sciences, Surgery, Umeå University, SE-901 87 Umeå, Sweden. Tel.: +46 736 260 800, fax.: +46 920 283 320. E-mail: lars.enochsson@umu.se (L. Enochsson).

Annals of Medicine & Surgery (2023) 85:2924–2931

Received 2 November 2022; Accepted 30 April 2023

Published online 15 May 2023

http://dx.doi.org/10.1097/MS9.0000000000000819

jaundice by inserting biliary stents[6]. This shift towards more invasive procedures further increases the demands on the endoscopist performing ERCP. It has been shown that endoscopists with less experience have higher rates of adverse events[7,8]. However, there are different opinions regarding the necessity of basic ERCP training since it is time-consuming and does not necessarily guarantee the skills required[9]. Furthermore, there are also diverging opinions on whether the training program should focus on a threshold number of procedures or be more focused at the individual's progress[10].

Small but well-designed prospective randomized studies show that simulator training of surgical residents in laparoscopic cholecystectomy improves the surgical performance[11–13]. However, reports in the literature on ERCP training on mechanical, hybrid, or virtual simulators are conflicting[14]. The development of advanced endoscopy simulators has paved the way for basic training as well as training in more complex endoscopic procedures such as ERCP. It is possible that simulator training in ERCP can reduce the frequency of adverse events and thus the risk for the patient. Furthermore, senior endoscopists may also benefit from simulator training by maintaining their ERCP skills[9].

Simulators for ERCP training include different types of simulators like mechanical, virtual, hybrid simulators, and *ex vivo* or *in vivo* training on animals. However, to our knowledge, there is no systematic review on the use of simulators in ERCP training. The primary aim of this systematic review was to study the extent to which simulators used for ERCP training are indeed properly validated. The secondary aim was to identify the role of simulator based ERCP training in clinical training.

This systematic review has been registered in the PROSPERO International prospective registry of systematic reviews 24/05/2022 https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022332614 [CRD42022332614].

Methods

This systematic review was conducted by searching the medical literature on MEDLINE and SCOPUS under the guidance of the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) protocol[15]. The search was conducted in February 2022 and included either observational, or randomized, or nonrandomized studies. The following criteria were required for inclusion of a study:

(1) English language.
(2) Abstract included.
(3) Published during the last 10 years in a peer-reviewed journal.

Search was performed by title or abstract, utilizing keywords, and Boolean operators as follows:

Medical Subject Heading (MeSH) terms and text words were used based on the following search strategy: (Endoscopic Retrograde Cholangiopancreatography) OR (ERCP); (simulation) OR (simulator); and #1 AND #2. The following filters were applied to #3: (Abstract) AND (in the last 10 years), AND (English), NOT (review).

Two of the authors (K.G. and L.E.) independently screened the abstracts of the articles that were considered to meet the inclusion criteria, and any differences in opinion were resolved by discussion. Then, a full-text review of those articles considered suitable for further examination was carried out.

Exclusion criteria were as follows:

(1) Reviews; meta-analyses. However, their reference lists of included papers were used to retrieve additional relevant studies ($n = 4$).
(2) Letters to the Editor.
(3) Nonpeer reviewed articles.
(4) Conference abstracts and proceedings.

Of the final papers from the search, only original studies conducted on endoscopists, surgeons, trainees, residents, and fellows of any endoscopy-related subspecialty were included in this systematic review. The flow-diagram of the selection process is shown in Figure 1.

The validity of each report was assessed according to the standards for educational and psychological testing. These standards state five main sources of evidence that can be used to support the validity of an interpretation for a new test or, in this case, to support the validity of a scoring system for a new simulator[16,17]. These five sources are: test content that is, the degree to which the test content corresponds to testing purposes; internal structure; response process; relationships to other variables (previously termed 'construct validity') that is, score correlation of a new assessment tool and an existing measure; and consequences of testing evidence.

Our method for identifying and evaluating data complied with the PRISMA 2020 statement and checklist[15], Supplemental

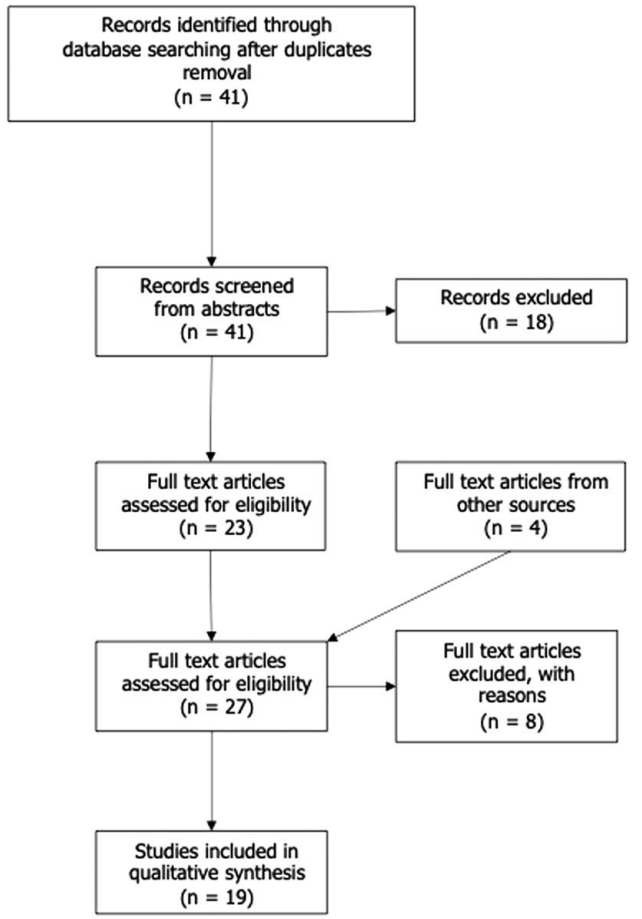

Figure 1. PRISMA flowchart of the selection process.

Digital Content 1, http://links.lww.com/MS9/A138, Supplemental Digital Content 2, http://links.lww.com/MS9/A139 and is reported in line with the assessment of methodological quality of systematic reviews (AMSTAR 2)[18], Supplemental Digital Content 3,. See Appendix A and the enclosed PRISMA Checklist, Supplemental Digital Content 1, http://links.lww.com/MS9/A138.

Results

Inclusion

Due to design differences of the reviewed articles, the Participants, Interventions, Comparators, and Outcomes tool could not be applied. Instead, we used the Sample, Phenomenon of Interest, Design, Evaluation, and Research type tool, which is based on Participants, Interventions, Comparators and Outcomes[19]. There was good compliance with the Amstar 2 tool, reporting 'Yes' for nine criteria, 'Partial Yes' for three, and 'No' for four, all related to meta-analyses and not applicable to the present review (Appendix A). The data of the articles included were extracted and checked independently using tables by two authors (K.G. and L.E.).

After duplicates were removed, the search strategy yielded 41 articles. Of these articles, 18 were excluded according to the predefined criteria through title and abstract screening. Twenty-three articles were selected for full-text review. In addition, four relevant articles were identified from the reference list of the reviewed papers and added to the study. The full-text review revealed eight articles that were excluded due to irrelevance. Ultimately, 19 papers were included in the study.

Characteristics of the articles included

The characteristics of the studies included are presented in Tables 1 and 2. Table 1 presents the 17 papers on the use of simulators in ERCP training, while Table 2 presents two papers comparing two training simulators. As shown in Table 1, new advanced endoscopy and ERCP simulators can be divided into mechanical (EMS)[14,20–24,29–32], *ex vivo*[26–28], hybrid[25,26], and digital[33–35] models. Five studies from Table 1[23,24,26,30,31] used a questionnaire or a Likert scale to evaluate the outcome, while most used cannulation rates and time spent to complete the procedure.

Risk for bias

To assess the risk for bias, we used the Newcastle–Ottawa quality assessment scale [https://www.ohri.ca//programs/clinical_epidemiology/oxford.asp]. The inspection of the studies under review revealed that not all domains of Newcastle–Ottawa could be addressed (Appendix B). Furthermore, bias was estimated regarding the missing data management as well as the type of funding, if any. It was impossible to apply the effect measure to the reviewed articles. The authors discussed all studies included in this review and any doubt about methodological quality was resolved and any limitations were further presented in the review. This comparison between most studies was impracticable due to the obvious heterogeneity of algorithms and datasets.

Validation

Some form of validation of the results was used in all but one of the studies[25].

The studies in Table 2 used a questionnaire to evaluate their findings[36,37].

Excluded studies

The 18 studies excluded and the reasons for their exclusion are given in Appendix C.

Main findings

In general, the EMS models were found to have high content validity and were effective for training cannulation, although the studies were small and lacked relevant control groups[14,20–24,29–32]. The experience from in vivo and ex vivo models was very limited, but they simulate the biological characteristics of the tissues adequately[26–28]. The digital models had poor content validity but were assumed to have the potential of technical improvement over time[33–35].

Discussion

ERCP training traditionally follows the apprenticeship model, whereby the novices observe and learn the various stages in clinical cases that gradually increase in complexity over time. Previous studies have shown that 350–400 ERCP procedures are needed to achieve a bile duct cannulation rate of 80%[38]. However, these numbers are usually impossible to achieve during a 3-year fellowship[39]. Furthermore, since these ERCP procedures are performed on patients, patient safety would be considerably improved if these 350–400 ERCP procedures were done by an experienced endoscopist. It is obvious that training on correctly validated ERCP simulators ought to play an important role during the early stages of a resident's career, allowing them to become familiar with the devices used and scope positioning before going on to clinical procedures. Several studies have been published where the purpose has been to validate simulators for ERCP or advanced endoscopy (Table 1). Regardless of whether it is a mechanical simulator[14,20–24,29–32], an *in vivo* or *ex vivo* model[27,28], hybrid simulator[25,26], or a virtual simulator[33–35], all assessments of tools in medical education require evidence of validity if they are to be interpreted in a correct and meaningful way.

As shown in Table 1, only one article fulfilled the criteria of a robust validation study[14] as stated by Downing[16]. However, it is important to stress that this simulator constructed by Jirapinyo *et al.*[14], although thoroughly validated, is a mechanical simulator that has its main place in the preclinical setting as it focus on the technical aspects of basic ERCP skills only.

Nearly all studies suffered from a small and heterogeneous sample size. In most studies, the use of ERCP simulators early on in training seems to accelerate the acquisition of practical skills. However, the use of learning curves is seldom applied. Most studies adequately assess response, but data on relationships to other variables and consequence of testing is more limited. Content validity also varies, as mechanical and virtual simulators aim at simulating different aspects of ERCP.

Programs for more advanced therapeutic techniques are usually not provided by a mechanical simulator. Thus, when training more advanced ERCP procedures, virtual simulators for endoscopy training such as Endo VR (CAE Healthcare)[33] or GI Mentor II (Surgical Science Sweden AB)[34,35] are used. The

Table 1

Characteristics of the reviewed articles.

References	Model's feature	Sample size	Validation	Outcome	Remarks
Lim*et al.* [20]	EMS	8 EMS 8 controls Multicenter randomized controlled trial.	Yes	Improved cannulation rates (69.6 vs 47.1%), lower total time (4.7 *vs* 10.3 mins); overall performance score not significantly different.	Short observation period, small number, and big variation of ERCPs performed by the trainees.
Liao*et al.* [21]	EMS	8 EMS 8 controls Two referral medical centers. Randomized controlled trial.	Yes	Improved deep biliary cannulation rates (73.3% *vs* 47.4%) and improved overall performance; No benefit of single *vs* multiple simulator practice.	Small number. Further studies are required to determine whether repeated coached practice can provide even greater benefit.
Meng*et al.* [22]	EMS	6 EMS 6 controls Randomized controlled trial.	Yes	Higher cannulation success in the EMS vs. control group ($P = 0.006$). Shorter mean times for intubation, cannulation, and completion ($P < 0.001$). EMS group had a significantly better mean performance score ($P = 0.006$).	Small number of participants. 150 ERCPs for each group. Difficult to discern to what extent coaching and simulator training had on the result. A large multicenter randomized controlled trial (RCT) is needed.
Frimberger*et al.* [23]	Frimberger's EMS for Billroth II or Roux-en-Y cannulation	10 trainees 11 experts	Yes A 7-point Likert scale	Trainees' skills improvement = 5.8, Experts: learning success of the trainees = 6.8. A simple and practicable simulator.	No training in negotiating the scope through the afferent limb to the papilla. A pilot study.
Schneider*et al.* [24]	EMS DIY*	10 trainees 6 experts	Impression questionnaire on a 7-point Likert scale.	Experts: Handling is close to reality 4.7, Training is reasonable for endoscopic education 5.3. Trainees: Skills improvement 6.4, 'I recommend For use in ERCP training' 6.4.	Not designed for advanced intraductal techniques but capable of improving the beginner's confidence.
Nguyen*et al.* [25]	Hybrid model	No	No	Reduces computation time up to 80%.	Just one *ex vivo* comparison.
Katanuma*et al.* [26]	A dry model for endoscopic sphincterotomy (ES) and needle knife precut sphincterotomy (NKP) using papilla of rolled noncured ham.	22	Questionnaire on a Likert scale from 1 to 10.	ES* was successful in 33 out of 34 attempts and NKP* was successful in all 7 attempts. Questionnaire: the median realism score was 7/10 for ES and 8/10 for NKP.	The model lacked information regarding bleeding and perforation. Further studies are needed to evaluate how this model contributes to the acquisition of ES and NKP skills.
Velázquez-Aviña*et al.* [27]	*Ex vivo* model using porcine stomach and chicken heart and trachea as papilla and bile ducts, respectively.	One experienced endoscopist	10 neo-papillae consecutively used in 1 duodenalized stomach.	In all cases: The stability of the neo-papilla was excellent. Cannulation, biliary sphincterotomy, and stent placement were successful.	Pilot study. Because the neo-papillae are interchangeable, repetitive interventions can be done using a single porcine model.
Itoi*et al.* [28]	*In vivo* and *ex vivo* porcine models using a hyaluronate solution simulated papilla.	*In vivo* model: 1 experienced endoscopist (EE) *Ex vivo* model: 1 EE and 2 trainees.	Validation of endoscopic sphincterotomy (ES) and endoscopic papillectomy (EP).	This novel porcine model appears useful for ES and EP training.	Pilot study. This current model does not allow cannulation or guidewire placement.
Voiosu*et al.* [29]	Boskoski-Costamagna EMS	Motion training group: $n = 16$. Standard group: $n = 20$ Multicenter, randomized controlled trial.	Yes	No significant difference of cannulation success between the two groups ($P = 0.37$). The motion training group had significantly lower median cannulation times and better technical performance on the first papilla type ($P = 0.013$).	Future studies are needed to establish Motion training in ERCP programs.
van der Wiel*et al.* [30]	Boskoski-Costamagna EMS	Beginners ($n = 11$) Intermediates ($n = 5$) Experienced ($n = 8$) Experts ($n = 22$).	Yes	Experts significantly faster than beginners. A high agreement among experts to include the simulator in the training of endoscopy beginners (3.86 on a 4-point Likert scale).	A good face and construct validity of the simulator.
van der Wiel*et al.* [31]	Boskoski-Costamagna EMS	40 ERCP experts in biliary sphincterotomy with synthetic papilla. Multicenter trial.	Yes	On a 10-point Likert scale: Realism = 7, Maneuvers resemblance = 8, Tactile feedback = 7, Realistic cutting = 6, and Cutting result = 8. A high agreement to include the papilla in the beginners training.	Good face validity.

Table 1

(*Continued*)

References	Model's feature	Sample size	Validation	Outcome	Remarks
Jirapinyo*et al.* [32]	Thompson Endoscopic Skills Trainer EMS.	42 participants including: 7 beginners, 7 1st-year, 7 2nd-year, 7 3rd-year GI fellows, 7 attending & 7 interventional attending physicians.	Yes	Aggregate scores differed significantly between training levels. Individual modules significantly differentiated between groups. Repetition improves score, with persistence of separation between training levels.	A part task box. 5 modules in a single session. It focuses only on the technical aspects and does not address the cognitive elements of endoscopic training.
Jirapinyo*et al.* [14]	Jirapinyo's EMS	23 participants: Beginners ($n=9$), Intermediates ($n=7$), Experienced ($n=7$).	Yes. Robust validation**	Excellent content validity indexes (CVIs) for all parameters measured.	A realistic, relevant representation of ERCP cannulation technique. It is effective at objectively assessing basic ERCP skills by differentiating scores based on clinical experience.
Sahakian*et al.* [33]	Endo VR	4 experienced 6 beginners	Total time to complete 4 simulated ERCP cases before and after 40 human ERCPs.	In baseline session, experts had a shorter procedure time than beginners. No significant difference in total time between experts and beginners between sessions 1 and 2.	The simulator was not responsive to an increase in trainee experience over time. Larger studies are needed to further evaluate the role of simulators in determining procedural competency.
Bittner*et al.* [34]	GI Mentor II 2 tasks (Case A and B)	3 1st-year and 3 2nd-year gastroenterology fellows, 3 gastroenterologists, and 3 GI surgeons.	Construct, content, and face validity. Didactic value assessed by questionnaire.	Mean procedure time defined skill levels. When outcomes of the two cases were combined together, beginners and experts differentiated based on time to complete the procedure, reach the papilla, and use fluoroscopy.	Other ERCP-specific metrics failed to demonstrate construct validity, likely secondary to the small sample size. Prospective, multicenter trials will be required to demonstrate the predictive validity of the GI Mentor II for ERCP.
Arnold*et al.* [35]	GI Mentor II and motion tracking	37 participants: 12 senior doctors who performed ERCP, 13 doctors with varying levels of experience, 12 untrained medical students.	1.Distance between the hands, 2.Height of the scope hand 3.Distance moved with the scope hand.	All motion tracking metrics discriminate between experts and beginners in specific sequences.	Further research needed

*DIY: Do-It-Yourself.
*ES = Endoscopic sphincterotomy.
*NKP = Needle knife precut sphincterotomy.
**See discussion in text draft.
EMS: ERCP mechanical simulator.

Table 2

Comparison articles.

References	Model comparison	Sample size	Validation	Outcome	Remarks
Leung *et al.* [36]	EMS *vs Ex vivo* Porcine Stomach Model (PSM).	22 experienced	Questionnaire	After practice, there was a significantly greater increase in confidence score for EMS than PSM ($P<0.003$). Participants found EMS more useful for training ($P=0.017$).	Subjective scores.
Leung *et al.* [37]	EMS vs GI Mentor II (ECS) **Digestive Endoscopy 2012	18 trainees 16 trainers	Questionnaires	Both EMS and ECS scored higher after practice. Both trainers & trainees showed significantly greater increases in scores for EMS vs. ECS in facilitating of the ERCP procedure, enhancing confidence in clinical ERCP. EMS scored significantly higher in realism and usefulness.	Because of time constraints the order of simulator practice was not randomized. The questionnaire evaluation is subjective by individuals with varying experience, even among the trainers.

advantages of these virtual simulators are that they have a modular database that covers many different situations, both diagnostic and therapeutic. The trainee can train sphincterotomy and extraction of bile duct stones, take brush cytology and biopsies, and insert pancreatic and bile duct stents.

However, despite all the sophisticated software and the well-developed face validity, in the studies validating virtual simulators in Table 1, only the mean procedure time could differentiate between novices and experts[33,34]. With motion tracking on the GI Mentor II simulator, better discrimination between novices and experts was achieved[35]. Despite the rather limited outcome regarding differences between beginners and experts, virtual simulators are too expensive for many hospitals and thus are only found in a limited number of larger hospitals and training centers. There are no widely accepted criteria for defining ERCP proficiency, although successful cannulation is usually considered a crucial moment[40,41] and the main focus of many ERCP simulation models.

In addition to simulated training, an organized curriculum and training program for ERCP should provide the base for acquiring the necessary skills while limiting the potential hazards of ERCP performed by inexperienced endoscopists. Such programs exist in some countries, but are lacking or poorly structured in others[42]. More facilities for ERCP training are necessary in both low and high-income countries.

ERCP training should preferably start in the early stages of residency if the goal is to accomplish proficiency as a specialist in gastroenterology or surgery. Although there are concerns about safety if a resident carries out an ERCP, these hazards may be overbridged if the training is started in a simulated environment. The schedule for training in advanced endoscopic techniques differs depending on the healthcare organization. Whereas residents may start performing ERCP during the last years of residency, the practice in the United States is to start the training during or after a gastroenterology fellowship.

Regardless of when the ERCP training is started, the endoscopist or surgeon should be closely supervised and be given continuous feedback after reaching satisfactory proficiency in a simulated model. Ideally, all procedures should be quality-assessed and registered in a database that enables continuous feedback on the success and safety of the procedures performed[43]. Clear and standardized criteria based on complication rate and successful canulation should be defined and applied before the endoscopist is allowed to carry out ERCP independently.

Conclusion

In this systematic review of validation studies of ERCP simulators, only one met the criteria for a correctly performed validation[14]. Furthermore, many of the studies reviewed were performed on a small number of participants and failed to evaluate the role of simulators in acquiring the skills necessary. Further studies are needed that include a sufficient number of participants, that evaluate a complete preclinical training program from beginner to competence to excellence on a simulator validated for ERCP, with continual assessment of competence throughout training.

Ethical approval

This is a systematic review of previous published articles. Ethical permit was thus not necessary.

Sources of funding

This work was supported by the Umeå County Council (ALF), Umeå University, Sweden.

Author contribution

K.G., K.T.A., and L.E.: conceived of the presented idea; K.G., K.T.A., and L.E.: designed the study; K.G. and L.E.: independently screened the abstracts and articles; K.G., K.T.A., G.S., and L.E. interpreted the data; K.G., N.O., and L.E. wrote the article; N.B. and G.S.: reviewed the article. All authors discussed the results and approved of the final manuscript.

Conflicts of interest disclosure

All authors declare that they have no conflicts of interest.

Research registration unique identifying number (UIN)

1. Name of the registry: PROSPERO International prospective register of systematic reviews.
2. Unique Identifying number or registration ID: CRD42022 332614.
3. Hyperlink to your specific registration (must be publicly accessible and will be checked): https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022332614.

Guarantor

Lars Enochsson, MD, PhD, is the guarantor of the submitted systematic review. He has full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Provenance and peer review

Not commissioned, externally peer-reviewed.

Acknowledgments

American Manuscript Editors and Peter Cox for language editing services.

References

[1] Finocchiaro M, Cortegoso Valdivia P, Hernansanz A, *et al*. Training simulators for gastrointestinal endoscopy: current and future perspectives. Cancers 2021;13:1427.

[2] Johnson G, Webster G, Boškoski I, *et al*. Curriculum for ERCP and endoscopic ultrasound training in Europe: european Society of Gastrointestinal Endoscopy (ESGE) position statement. Endoscopy 2021; 53:1071–87.

[3] Dumonceau J-M, Andriulli A, Elmunzer BJ, *et al*. European society of gastrointestinal endoscopy, prophylaxis of post-ERCP pancreatitis: european Society of Gastrointestinal Endoscopy (ESGE) guideline updated June 2014. Endoscopy 2014;46:799–815.

[4] Österberg J, Sandblom G, Enochsson L, Annual report 2021 - GallRiks, (2021). https://www.ucr.uu.se/gallriks/fou/arsrapporter/arsrapporter/ars rapport-2020.

[5] Ding X, Zhang F, Wang Y. Risk factors for post-ERCP pancreatitis: a systematic review and meta-analysis. Surgeon 2015;13:218–29.

[6] Freeman ML. Complications of endoscopic retrograde cholangiopancreatography: avoidance and management. Gastrointest Endosc Clin N Am 2012;22:567–86.

[7] Lee HJ, Cho CM, Heo J, *et al*. Impact of hospital volume and the experience of endoscopist on adverse events related to endoscopic retrograde cholangiopancreatography: a prospective observational study. Gut Liver 2020;14:257–64.

[8] Syrén E-L, Sandblom G, Enochsson L, *et al*. Outcome of ERCP related to case-volume. Surg Endosc 2022;36:5339–47.

[9] Gallo C, Boškoski I, Matteo MV, *et al*. Training in endoscopic retrograde cholangio-pancreatography: a critical assessment of the broad scenario of training programs and models. Expert Rev Gastroenterol Hepatol 2021; 15:675–88.

[10] Ekkelenkamp VE, Koch AD, de Man RA, *et al*. Training and competence assessment in GI endoscopy: a systematic review. Gut 2016; 65:607–15.

[11] Seymour NE, Gallagher AG, Roman SA, *et al*. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Ann Surg 2002;236:458–63; discussion 463-4.

[12] Grantcharov TP, Kristiansen VB, Bendix J, *et al*. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. Br J Surg 2004;91:146–50.

[13] Ahlberg G, Enochsson L, Gallagher AG, *et al*. Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. Am J Surg 2007;193: 797–804.

[14] Jirapinyo P, Thompson AC, Aihara H, *et al*. Validation of a novel endoscopic retrograde cholangiopancreatography cannulation simulator. Clin Endosc 2020;53:346–54.

[15] Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Int J Surg 2021;88: 105906.

[16] Downing SM. Validity: on meaningful interpretation of assessment data. Med Educ 2003;37:830–7.

[17] American Educational Research Association, American Psychological Association, National Council on Measurement in Education, The Standards for Educational and Psychological Testinga, American Psychological Association, 2014.

[18] Shea BJ, Reeves BC, Wells G, *et al*. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017;358:j4008.

[19] Cooke A, Smith D, Booth A. Beyond PICO: The SPIDER tool for qualitative evidence synthesis. Qual Health Res 2012;22:1435–43.

[20] Lim BS, Leung JW, Lee J, *et al*. Effect of ERCP mechanical simulator (EMS) practice on trainees' ERCP performance in the early learning period: US multicenter randomized controlled trial. Am J Gastroenterol 2011;106:300–6.

[21] Liao W-C, Leung JW, Wang H-P, *et al*. Coached practice using ERCP mechanical simulator improves trainees' ERCP performance: a randomized controlled trial. Endoscopy 2013;45:799–805.

[22] Meng W, Yue P, Leung JW, *et al*. Impact of mechanical simulator practice on clinical ERCP performance by novice surgical trainees: a randomized controlled trial. Endoscopy 2020;52:1004–13.

[23] Frimberger E, Abdelhafez M, Schmid RM, *et al*. A novel mechanical simulator for cannulation and sphincterotomy after Billroth II or Roux-en-Y reconstruction. Endosc Int Open 2016;4:E922–6.

[24] Schneider AR, Schepp W. Do it yourself: building an ERCP training system within 30 minutes (with videos). Gastrointest Endosc 2014;79: 828–32.

[25] Nguyen T, Kim M, Lee D. A hybrid contact model for cannulation simulation of ERCP. Stud Health Technol Inform 2014;196:304–6.

[26] Katanuma A, Itoi T, Umeda J, *et al*. A novel dry model for practicable sphincterotomy and precut needle knife sphincterotomy. Gastroenterol Res Pract 2014;2014:908693.

[27] Velázquez-Aviña J, Sobrino-Cossío S, Chávez-Vargas C, *et al*. Development of a novel and simple ex vivo biologic ERCP training model. Gastrointest Endosc 2014;80:1161–7.

[28] Itoi T, Gotoda T, Baron TH, *et al*. Creation of simulated papillae for endoscopic sphincterotomy and papillectomy training by using in vivo and ex vivo pig model (with videos). Gastrointest Endosc 2013;77:793–800.
[29] Voiosu T, Puscasu C, Orlandini B, *et al*. Motion training on a validated mechanical ERCP simulator improves novice endoscopist performance of selective cannulation: a multicenter trial. Endosc Int Open 2021;9:E145–51.
[30] van der Wiel SE, Koch AD, Bruno MJ. Face and construct validity of a novel mechanical ERCP simulator. Endosc Int Open 2018;6:E758–65.
[31] van der Wiel SE, Koch AD, Bruno MJ. Face validity of a synthetic papilla designed for biliary sphincterotomy training. Endosc Int Open 2019;7: E757–61.
[32] Jirapinyo P, Kumar N, Thompson CC. Validation of an endoscopic part-task training box as a skill assessment tool. Gastrointest Endosc 2015;81: 967–73.
[33] Sahakian AB, Laine L, Jamidar PA, *et al*. Can a computerized simulator assess skill level and improvement in performance of ERCP? Dig Dis Sci 2015;61:722–30.
[34] Bittner JG, Mellinger JD, Imam T, *et al*. Face and construct validity of a computer-based virtual reality simulator for ERCP. Gastrointest Endosc 2010;71:357–64.
[35] Arnold SH, Svendsen MB, Konge L, *et al*. Three-dimensional motion tracking correlates with skill level in upper gastrointestinal endoscopy. Endoscopy 2015;47:825–8.
[36] Leung JW, Wang D, Hu B, *et al*. A head-to-head hands-on comparison of ERCP mechanical simulator (EMS) and Ex-vivo Porcine Stomach Model (PSM). J Interv Gastroenterol 2011;1:108–13.
[37] Leung J, Lim B, Ngo C, *et al*. Head-to-head comparison of practice with endoscopic retrograde cholangiopancreatography computer and mechanical simulators by experienced endoscopists and trainees. Dig Endosc 2012;24:175–81.
[38] Verma D, Gostout CJ, Petersen BT, *et al*. Establishing a true assessment of endoscopic competence in ERCP during training and beyond: a single-operator learning curve for deep biliary cannulation in patients with native papillary anatomy. Gastrointest Endosc 2007;65:394–400.
[39] Kowalski T, Kanchana T, Pungpapong S. Perceptions of gastroenterology fellows regarding ERCP competency and training. Gastrointest Endosc 2003;58:345–9.
[40] Katzarov AK, Dunkov ZI, Popadiin I, *et al*. How to measure quality in endoscopic retrograde cholangiopancreatography (ERCP). Ann Transl Med 2018;6:265.
[41] Elmunzer BJ, Walsh CM, Guiton G, *et al*. SVI study group and the U.S. Cooperative for Outcomes Research in Endoscopy. Development and initial validation of an instrument for video-based assessment of technical skill in ERCP. Gastrointest Endosc 2021;93:914–23.
[42] Teoh AYB, Ma MTW, Aabakken L, *et al*. International cross-sectional survey on endoscopic retrograde cholangiopancreatography training by the World Endoscopy Organization. Dig Endosc 2022;35: 103–110.
[43] Enochsson L, Thulin A, Österberg J, *et al*. The swedish registry of gallstone surgery and endoscopic retrograde cholangiopancreatography (GallRiks): a nationwide registry for quality assurance of gallstone surgery. JAMA Surgery 2013;148:471–8.