Introduction

ERCP is considered to be an advanced endoscopic procedure and is one of the most technically demanding and high-risk procedures in gastrointestinal endoscopy. The procedure may result in life-threatening short-term and long-term complications, and therefore, ERCP carries a higher risk for morbidity and mortality than most other endoscopic procedures [1]. Achieving competence in ERCP requires a great deal of training and an extensive number of procedures, but there is still no consensus on assessment of a trainee’s competence. Currently, in most training centers around the world, threshold numbers are still used as a surrogate for competency and certification. The first publications concerning the minimal number of ERCP procedures that must have been completed to gain technical competence were published back in the early 1990s and report numbers ranging from 100 to 180 procedures [2–4]. The current guideline of the American Society for Gastrointestinal Endoscopy (ASGE) recommends mastery of at least 200 ERCP procedures to gain competence [5]. However, a recent study by Cotton et al. describes that only a minimal number of hospitals in the United States adhere to these guidelines [6]. Therefore, there is growing evidence that competence should be established by objective performance criteria [7–9].
Nowadays, trainees learn to perform ERCP with hands-on training in a clinical setting on real patients under the supervision of a trained endoscopist: the so-called master apprentice model. This setting offers immediate feedback from an experienced endoscopist, but it also has certain drawbacks. Trainees learn by “trial and error,” which potentially increases patient discomfort and risk of complications and prolongs procedure time, which has additional economic consequences. Simulators offer the potential to train in a dedicated “learning environment, offering a less stressful situation with less potential risks and the opportunity to endlessly repeat specific tasks.

Simulator-based gastrointestinal endoscopy training has gained acceptance over recent decades and has been extensively studied. Multiple simulators have been validated and it has been demonstrated that use of simulators in gastrointestinal endoscopy accelerates the early learning curve of trainees [10–13]. The proven value of simulator-based training has led to the recommendation to introduce gastrointestinal simulators in the curriculum for endoscopists being trained in forward-viewing endoscopy [12]. Simulator-based training seems ideally suited as a training platform for ERCP, due to the complexity of this procedure and its associated complications. Although simulator training in ERCP has been possible for quite some time now, there are very limited scientific data on application of endoscopic simulators in ERCP in training novice endoscopists. Six simulators have been described in the literature. Nevertheless, despite their definite training potential, the applicability of each of the simulators as a certified training tool has not been demonstrated. Before using a novel simulator as a training tool, it is essential to carry out a study to scientifically determine its validity. Studies evaluating face and construct validity demonstrate the appropriateness of using a simulator for training or assessment [14, 15]. Validity assessment of a medical simulator can be performed on different levels. One of the most commonly used forms of validation is face validity, in which a defined group of subjects are asked to judge the degree of resemblance between a training simulator and the real activity. This is often combined with construct validity. This describes the degree to which the assessment can discriminate between different experience levels. The most powerful evidence is gained through concurrent validity, in which performance on the system is compared with outcomes from an established assessment method designed to measure the same skills or attributes. This implies that experience gained by training on the simulator results in improved performance in patient-based procedures [13–15].

The Boškoski-Costamagna ERCP Trainer is a novel mechanical ERCP training model developed and produced by Cook Medical (Cook Medical, Limerick, Ireland) in close collaboration with the Digestive Endoscopy Unit of the Gemelli Hospital in Rome, Italy (Dr. Boškoski and Prof. Dr. Costamagna). It is designed to guide trainees on how to correctly position the endoscope in front of the papilla in order to attain a proper axis and to achieve deep cannulation from where several therapeutic interventions can be performed, such as plastic or metal stent placement and stone extraction [16]. An initial report has been published by Jovanovic et al showing the potential value of the model [17]. However, no attempt was made to scientifically assess the validity of the model. The current study had three aims: to determine whether the simulator can distinguish between endoscopists with different levels of experience (construct validity); to evaluate the extent to which the ERCP simulator simulates actual ERCP procedures (face validity); and to assess the value of the simulator as a training tool, as judged by experts.

**Methods**

**Simulator**

The Boškoski-Costamagna ERCP Trainer (Cook Medical, Limerick, Ireland) is a mechanical endoscopic simulator (Fig.1 and Fig. 2). The simulator has been designed to train residents on correct positioning of the endoscope, based on the knowledge that a successful ERCP procedure is largely dependent upon the ability to achieve an optimal position of the scope tip in front of the papilla. The model consists of a metal framework with the esophagus, stomach, and duodenum constructed from plastic. The various papillae are made out of latex with both bile and pancreatic ducts inserted in different varieties intended to resemble known anatomical variations. The simulator enables use of a real endoscope and commonly used equipment. The model can be placed on a table and real-time fluoroscopy image is made visible on a secondary screen using a small camera (Fig. 3). The Boškoski-Costamagna ERCP Trainer can simulate different patient positions (prone, oblique and supine) and a variety of ERCP procedures: scope insertion, wheel and elevator handling, selectively cannulating the bile and pancreatic duct, stone extraction, and both metal and plastic stent insertion. The level of difficulty can be adjusted based on variations of the papillary anatomy and biliopancreatic junction [16]. The version using during this entire study was the second generation of the ERCP Trainer.

**Participants**

We included four groups of participants in this study: novices, intermediates, experienced, and experts. The participants were divided into these groups based on lifetime ERCP experience. There is no consensus in the literature when it comes to numbers expressing experience levels in ERCP. Therefore, we attempted to define the groups according to the largest reported numbers in the literature [18–20]. The bar for the expert group was deliberately raised to 2500 ERCPs lifetime to reassure an expert group with an undisputed reputation. The first group, the novices, was defined as participants with less than 50 lifetime ERCPs. The second group, the intermediates, had a lifetime experience of 50 to 600 ERCPs. The third group consisted of experienced participants with a lifetime experience of 601 to 2500 ERCPs. Based on previous studies concerning ERCP, we assumed that experts would be twice as fast at completing the ERCP simulator assignments compared to novices [19, 21, 22]. A sample size calculation reveals a minimal sample of seven participants both in the novice and expert groups to achieve a power of 0.80.
All participants were invited for a simulator session in a similar private conference room either in our hospital, during a national conference of the Dutch Society of Gastroenterology in spring 2016 or during Digestive Disease Week in May 2016. During the procedures the participants were assisted by five alternating endoscopy nurses from the Erasmus MC with broad experience in assisting ERCP procedures and good English conversation skills.

Questionnaire
All participants filled out a questionnaire on demographics, medical experience, and endoscopy experience. Endoscopy experience included the number of various endoscopic procedures performed annually and estimated lifetime numbers. In addition, participants were asked about previous experience with other medical simulators. After performing a standardized set of assignments on the simulator, experts were asked to rate their appreciation of the realism of the ERCP Trainer. Appreciation was expressed on a 10-point Likert scale [23], varying from very unrealistic (1) to very realistic (10). Questions were asked about the realism of the simulator setup, anatomical representation, difficulty, handling of the endoscope, haptic feedback, and imaging. In addition, experts were asked to evaluate the didactic value of the Boškoski-Costamagna ERCP Trainer on a four-point Likert Scale.

ERCP simulation
All participants were invited to perform six standardized assignments. The first assignment was to establish the correct position of the endoscope in front of the papilla to gain a proper axis for cannulation. The second assignment was to cannulate the common bile duct (CBD) during which the number of unintentional pancreatic duct (PD) cannulations was also scored. The same applied for PD cannulation and unintentional CBD cannulations. Next, participants were asked to place a plastic stent in the PD and a plastic stent in the CBD. Finally, participants were asked to extract a single stone from the CBD using an extraction basket. A coffee bean was used as a stone. For each exercise, time was recorded, with a time limit of 10 minutes per assignment for logistical reasons. After 10 minutes the procedure was scored as failed. Participants were not made aware of the time limit at the start of the assignment and they were encouraged to complete the exercises to best of their ability without a competitive intent.
Data analysis

All statistical analyses were performed using SPSS 21.0 software (IBM Corp: Armonk, NY). Descriptive statistics were used for all measures. Assuming that the experts required less time to fulfill the assignments compared to novices, variations in outcomes between groups were compared using a Kruskal-Wallis test. A separate analysis between the four groups was performed using a Mann-Whitney U Test. Data are presented as median and range. A P value < 0.05 was considered significant.

Results

Participants

In total, 46 participants were included in the study, 11 novices, 5 intermediates, 8 experienced ERCPists and 22 ERCP experts, they originated from 20 different countries from all continents. The percentage of female participants was 21.7% and all ERCP experts were male. The mean number of years of endoscopic experience was 23.32 (range 15–34) for experts and 1.6 years for novices (range 0.1 – 5.0). Novices claimed 100% familiarity with previous use of endoscopy simulators, experts reported 77.3%. All participants completed the exercises and filled out the evaluation form. ▶ Fig. 4 shows the study design and baseline characteristics can be found in ▶ Table 1.

Face validity

The ERCP experts rated the Boškowski-Costamagna ERCP Trainer 7.12 on a 10-point Likert scale for overall realism. ▶ Table 2 details the experts’ average ratings of various components of the simulator. ▶ Table 3 demonstrates the perceived opinion of the ERCP-specific components, as judged by experts. In general, most of the experts rated the more complex procedural aspects of the Boškowski-Costamagna ERCP Trainer as very realistic. Biliary plastic stent placement scored 7.99 on a 10-point Likert scale, pancreatic plastic stent placement 7.80 and removal of a common bile duct stone 7.42.

Construct validity

Construct validity of the Boškowski-Costamagna ERCP Trainer was evaluated by comparing the procedure time and attempts at success per assignment across the four groups. Data output regarding the construct validity is presented in ▶ Table 4 and ▶ Table 5. All assignments were completed faster by the experts (P = 0.000), experienced (P = 0.000) and intermediate (P = 0.052) than by the novices (▶ Table 3). Novices (n = 11) completed the procedure in a mean time of 21.09 (min:sec), intermediates (n = 5) in 10.58, experienced (n = 8) in 6.42, and experts (n = 22) in 6.05. Unintentional CBD cannulation occurred with a median of 4.6 in novices, 0.8 for intermediates, 2.0 for experienced, and 1.4 unintended CBD cannulations for experts (P = 0.028). Unintended PD cannulation occurred less frequently, median of 0.4 for novices, 0 unintended cannulations for intermediates and experienced, and 0.2 for experts (P = 0.449). There were no statistical differences between novices and experts in the number of attempts per assignment (P = 0.985). We performed a separate analysis between the four groups using a Mann-Whitney U test. No statistical significant differences were seen between the intermediate, experienced and ex-
erts, all of whom performed all assignments faster than the novices (▶Table 4).

Didactic/training value
Expert opinion was that the Boškoski-Costamagna ERCP Trainer is a useful tool in the basic training of a novice endoscopist (3.91 on a 4-point Likert scale) and that the ERCP trainer should be incorporated into the training of novice endoscopists (3.86 on a 4-point scale). The expertise gained on this simulator should be applicable in a clinical curriculum (3.59 on a 4-point scale). Most experts agreed that there is a limited role for the simulator in training of experienced ERCP-performing endoscopists (rated 1.86 on a 4-point scale).

Discussion
This study reports the first formal validation of the Boškoski-Costamagna mechanical ERCP Trainer and demonstrates good construct and face validity. We demonstrated results of endoscopists who originated from all over the world, classified in four different expertise levels based on ERCP lifetime experience. The data reveal that the simulator is able to discriminate between different levels of expertise. Both experienced and expert endoscopists demonstrated superior performance, compared to novices, on all parts of the ERCP procedure.

All experts agree that the Boškoski-Costamagna ERCP Trainer seems to offer a satisfactory representation of clinical ERCP and that the expertise gained on the simulator should be transferrable to a clinical curriculum. Tactile feedback from the simulator was evaluated positively, even though the mechanical ERCP simulator is constructed from plastic and rubber components. The specific strengths of the simulator are the high levels of realism of more complex ERCP interventions, such as stent placement and stone extraction. This means that this novel ERCP simulator provides a platform with which to train inexperienced endoscopists in these complex procedures until they feel comfortable and perform up to a certain standard before exposing them to actual patient procedures. Our data do not support the use of this simulator in training endoscopists who already have a more experienced performance level.

A recent systematic review by our study group [10] presented an overview of currently available simulators and their known potential in training novices. Only six simulators have previously been described in the literature [18, 19, 21, 24–26]. A comparison validation study of three of the available simulators was performed by Sedlack et al. [27]. They included the Erlangen Endo Trainer, a bio simulation model, the live porcine model and the GI Mentor II, a virtual reality simulator. The study describes the potential value of all the simulators in training novice endoscopists, but they all have certain advantages and disadvantages. In terms of realism, bio-simulation models and live porcine models scored higher than mechanical and virtual reality simulators. However, major drawbacks of these types of simulators include costs and organizational difficulties due to the ethical incidentals. The X-Vision ERCP training simulator and the ERCP mechanical simulator have not been included in the study by Sedlack et al., both mechanical simulators [21, 22]. The X-vision ERCP trainer has been validated, but no studies have been published on implementation of the model in a training setting. The ERCP Mechanical Simulator has proven its training value in novice trainees [28].

There were some limitations in our study. It would have been ideal to include only naive trainees in the group with novices, whereas we included five novices without any experience and six novices who had performed fewer than 50 ERCPs. There is bias in terms of exposure and experience. However, with the
current ASGE guidelines in mind, defining ERCP competence after at least 200 cases [5], our novices are all in their very early learning curve. Another limitation might be use of procedure time as a surrogate marker for outcome. Time as a surrogate marker for outcome is not ideal. However, in many simulation validation studies, use of procedure times is unavoidable. This is an accepted method where participants are not made aware of the time element and encouraged not to give their fastest but to give their best performance [21, 29–32].

We believe that, compared to the previously described simulators, the Boškoski-Costamagna ERCP Trainer has added value. It has the advantage that a real endoscope and real accessories are used, providing novice endoscopists with the opportunity to learn how to handle the movements of the endoscope and experience the haptic feedback of their actions. A fluoroscopic image is created with the use of a simple camera, without the need for specific x-ray equipment. The total set-up creates the feel of a standard endoscopy unit. The Boškoski-Costamagna ERCP Trainer is light enough to be transported and easy to set up, and its use is not restricted to a specific environment. Despite the mechanical aspect of the model, its realism was scored satisfactory by experts. Experts believe that the Boškoski-Costamagna ERCP Trainer will improve trainee performance in the early training setting.

**Conclusion**

The Boškoski-Costamagna ERCP Trainer demonstrates good face and construct validity as a novel simulator for basic ERCP training. Experts generally agree on the didactic strength and added value of this simulator in the training curriculum for novice endoscopists.

**Acknowledgements**

This study was funded by an unrestricted grant for investigator-initiated research by Cook Medical, Limerick, Ireland.

---

**Table 4** ERCP assignments overview in procedural time.

<table>
<thead>
<tr>
<th>Procedure</th>
<th>Novice n=11</th>
<th>Intermediate n=5</th>
<th>Experienced n=8</th>
<th>Expert n=22</th>
<th>P value&lt;sup&gt;1&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positioning endoscope</td>
<td>Mean 01:07</td>
<td>00:26</td>
<td>00:11</td>
<td>00:14</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Median 00:59</td>
<td>00:15</td>
<td>00:10</td>
<td>00:12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Range 00:25–02:31</td>
<td>00:09–01:08</td>
<td>00:08–00:20</td>
<td>00:06–00:40</td>
<td></td>
</tr>
<tr>
<td>CBD cannulation</td>
<td>Mean 02:03</td>
<td>01:36</td>
<td>00:48</td>
<td>00:39</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Median 02:05</td>
<td>00:52</td>
<td>00:45</td>
<td>00:40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Range 00:37–04:05</td>
<td>00:28–03:24</td>
<td>00:11–01:14</td>
<td>00:05–01:31</td>
<td></td>
</tr>
<tr>
<td>PD cannulation</td>
<td>Mean 05:47</td>
<td>01:55</td>
<td>01:09</td>
<td>01:06</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>Median 06:27</td>
<td>01:17</td>
<td>01:13</td>
<td>00:27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Range 00:12–10:00</td>
<td>00:13–06:18</td>
<td>00:05–03:58</td>
<td>00:06–05:17</td>
<td></td>
</tr>
<tr>
<td>PD stent placement</td>
<td>Mean 03:35</td>
<td>00:43</td>
<td>00:45</td>
<td>00:39</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Median 06:53</td>
<td>00:47</td>
<td>00:48</td>
<td>00:38</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Range 01:15–10:00</td>
<td>00:30–00:49</td>
<td>00:26–01:12</td>
<td>00:13–01:27</td>
<td></td>
</tr>
<tr>
<td>CBD stone extraction</td>
<td>Mean 05:11</td>
<td>03:05</td>
<td>02:31</td>
<td>01:25</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>Median 03:55</td>
<td>01:46</td>
<td>02:41</td>
<td>01:06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Range 01:38–10:00</td>
<td>00:47–10:00</td>
<td>00:32–04:24</td>
<td>00:37–03:32</td>
<td></td>
</tr>
<tr>
<td>CBD stent placement</td>
<td>Mean 04:21</td>
<td>03:11</td>
<td>01:15</td>
<td>01:25</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Median 03:42</td>
<td>01:56</td>
<td>01:08</td>
<td>01:06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Range 01:14–10:00</td>
<td>01:03–08:38</td>
<td>00:51–01:48</td>
<td>00:37–03:32</td>
<td></td>
</tr>
<tr>
<td>Total procedure time</td>
<td>Mean 21:09</td>
<td>10:58</td>
<td>06:42</td>
<td>06:05</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Median 20:21</td>
<td>10:20</td>
<td>06:33</td>
<td>05:39</td>
<td></td>
</tr>
</tbody>
</table>

Procedural time in mm:ss
CBD, common bile duct; PD, pancreatic duct
<sup>1</sup>Kruskal-Wallis Test
Participants
Ang Tiing Leong, MD, PhD
Antonie J. P. van Tilburg, MD, PhD
Arjun D. Koch, MD, PhD
Bart Opsteeg, MD
Benedict Devereaux, MD, PhD
David F. Boerwinkel, MD, PhD
Dorien J. Buurman, MD, PhD
Dorien Oude Hergelink, MD, PhD
Ernst J. Kuipers, MD, PhD
Fauze Maluf-Filho, MD, PhD
Femme Harinck, MD, PhD
Fernando Marson, MD, PhD
Francisco M. Tufare, MD, PhD
Frank Ter Borg, MD, PhD
Fred M. Konikoff, MD, PhD
George J. Webstern, MD, PhD
Hrvoje Ivhecovic, MD, PhD
Hugo Richter Roca, MD, PhD
Inge van Rongen, MD, PhD
Jesús Rodriguez, MD, PhD
Jorge Cerecedo, MD, PhD
Judith E. Baars, MD, PhD
Lars Aabakken, MD, PhD
Lars Enochsson, MD, PhD
Lieke Hol, MD, PhD
Marc Giovanni, MD, PhD
Marjan R. S. Drenth, MD
Martin Goetz, MD, PhD
Martin W. James, MD, PhD
Matthijs P. Schwartz, MD, PhD
Michael L. Kochman, MD, PhD
Milutin Bulajic, MD, PhD
Moniek H. P. Gorter, MD
Paola Figueroa-Barojas, MD, PhD
Paul Didden, MD
Pietro Familiari, MD, PhD
Rastislav Kunda, MD, PhD
Raul F. Matano, MD, PhD
Renato Baracat, MD, PhD
Sarah N. Mutsaers, MD
Stefan Seewald, MD, PhD
Takai Itoi, MD, PhD
Thierry Ponchon, MD, PhD
Todd H. Baron, MD, PhD
Wout O. A. Rohof, MD, PhD

Competing interests
Dr. van der Wiel has nothing to disclose. Dr. Bruno reports institutional study support by unrestricted grant from Cook Medical, Consultancy and lecturer Fees from Boston Scientific and Cook Medical. Dr. Koch reports Consultancy and lecturer Fee from Cook Medical.

References

Table 5 Differences between groups.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Positioning endoscope</td>
<td>Chi-square</td>
<td>6.500</td>
<td>0.000</td>
<td>6.000</td>
<td>10.000</td>
</tr>
<tr>
<td></td>
<td>Asymp.sign.</td>
<td>0.013</td>
<td>0.000</td>
<td>0.171</td>
<td>0.186</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBD Cannulation</td>
<td>Chi-square</td>
<td>17.000</td>
<td>0.007</td>
<td>17.000</td>
<td>12.000</td>
</tr>
<tr>
<td></td>
<td>Asymp.sign.</td>
<td>0.267</td>
<td>0.000</td>
<td>0.284</td>
<td>0.064</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PD Cannulation</td>
<td>Chi-square</td>
<td>12.000</td>
<td>0.016</td>
<td>42.000</td>
<td>14.000</td>
</tr>
<tr>
<td></td>
<td>Asymp.sign.</td>
<td>0.090</td>
<td>0.002</td>
<td>0.435</td>
<td>0.447</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PD stent placement</td>
<td>Chi-square</td>
<td>0.000</td>
<td>0.000</td>
<td>2.500</td>
<td>17.500</td>
</tr>
<tr>
<td></td>
<td>Asymp.sign.</td>
<td>0.002</td>
<td>0.000</td>
<td>0.724</td>
<td>0.485</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBD Stone extraction</td>
<td>Chi-square</td>
<td>11.500</td>
<td>0.152</td>
<td>50.000</td>
<td>17.000</td>
</tr>
<tr>
<td></td>
<td>Asymp.sign.</td>
<td>0.069</td>
<td>0.006</td>
<td>0.724</td>
<td>0.976</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBD Stent placement</td>
<td>Chi-square</td>
<td>18.000</td>
<td>0.000</td>
<td>22.000</td>
<td>8.000</td>
</tr>
<tr>
<td></td>
<td>Asymp.sign.</td>
<td>0.320</td>
<td>0.000</td>
<td>0.093</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total procedure time</td>
<td>Chi-square</td>
<td>10.000</td>
<td>0.052</td>
<td>6.000</td>
<td>18.000</td>
</tr>
<tr>
<td></td>
<td>Asymp.sign.</td>
<td>0.000</td>
<td>0.000</td>
<td>0.833</td>
<td>0.377</td>
</tr>
</tbody>
</table>

Mann-Whitney U Test, two-tailed test, exact significance
CBD, common bile duct; PD, pancreatic duct


