Thieme

Impact of trainee involvement on the outcome of ERCP procedures: results of a prospective multicenter observational trial

Authors

Theodor Voiosu[1,2], Ivo Boskoski[3,4], Andrei M. Voiosu[1], Andreea Benguș[1], Agata Ladic[5], Ivo Klarin[6], Vincenzo Bove[3,4], Bogdan Busuioc[7], Mihai Rimbaș[1,2], Nadan Rustemovic[5], Bogdan Mateescu[1,2], Ivan Jovanovic[8], Guido Costamagna[3,4]

Institutions

1 Gastroenterology Department, Colentina Clinical Hospital, Bucharest, Romania
2 Internal Medicine Department, Carol Davila School of Medicine, Bucharest, Romania
3 Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy
4 Università Cattolica del Sacro Cuore, Centre for Endoscopic Research Therapeutics and Training (CERTT), Rome, Italy
5 University Hospital Centre, Zagreb, Croatia
6 Gastroenterology Department, Zadar General Hospital, Zadar, Croatia
7 Endoscopy Division, Cantacuzino Clinical Hospital, Bucharest, Romania
8 Clinic for Gastroenterology and Hepatology, University of Belgrade Medical School, Belgrade, Serbia

submitted 22.7.2019
accepted after revision 25.10.2019

Bibliography
DOI https://doi.org/10.1055/a-1049-0359
Published online: 25.11.2019 | Endoscopy 2020; 52: 115–122

ISSN 0013-726X

Corresponding author
Theodor Voiosu, MD, Gastroenterology Department, Colentina Clinical Hospital, 19-21 Stefan cel Mare BVD, Bucharest, Romania
Fax: +40-21-3180604
theodor.voiosu@gmail.com

Appendix 1s
Online content viewable at:
https://doi.org/10.1055/a-1049-0359

Scan this QR-Code for the author commentary.

ABSTRACT

Background Training in advanced endoscopic techniques such as endoscopic retrograde cholangiopancreatography (ERCP) should be driven by key performance measures and standardized competence assessment in order to provide safe and high-quality interventions. We aimed to determine whether the involvement of trainees influences the outcome of the procedure and the incidence of ERCP-related adverse events.

Methods This was an international, multicenter, prospective, observational study conducted at six high- and low-volume centers across Europe between October 2016 and October 2018, and included independent operators and their trainees. Standard report forms documenting indication, trainee involvement, technical outcome, and complications over a 30-day follow-up of consecutive ERCP procedures were included in the analysis. Technical success of the procedure and procedure-related adverse events were compared between procedures in the trainee group and the control group using bivariable and multivariable analysis.

Results 21 trainees and 16 control endoscopists performed 1843 ERCPs during the study period. Trainee involvement in ERCP procedures did not decrease technical success (92.4% vs. 93.7%; $P=0.30$) or increase the risk of adverse events (14.7% vs. 14.6%; $P>0.99$). Conversely, there were significantly more moderate or severe adverse events in the control group compared with the trainee group (6.2% vs. 3.4%, $P=0.01$). On multivariable analysis, only increased bilirubin levels, time to cannulation, and procedure difficulty level increased the risk of any procedure-related adverse event.

Conclusion Trainee involvement in ERCP interventions within a proper teaching setting is safe and does not compromise the success of the procedure.

Introduction

Training for advanced endoscopic procedures such as endoscopic retrograde cholangiopancreatography (ERCP) remains an important challenge for most training programs across the world. In order to address this issue, standards for training duration, procedure volume, and competence assessment have been proposed by several societies [1], and key performance measures for independent ERCP practice have been defined [2]. However, it has been widely recognized that many endoscopists in training fail to reach the required procedure volume or the predefined competence threshold during their training period [3, 4]. To compound this issue, there are relatively limited and conflicting data pertaining to the safety of ERCP procedures with trainee involvement [5, 6]. We hypothesize that in a teaching setting, trainee participation in ERCP procedures is safe and does not compromise the technical success of the procedure. This is very important because, as the incidence and severity of adverse events during ERCP, especially in low-volume centers, is not negligible, assuring patient safety remains of paramount importance.

In our study, we aimed to prove that ERCP procedures with trainee involvement do not differ significantly from procedures completed without trainee involvement, both in terms of technical success and procedure-related adverse events.

Methods

We conducted an investigator-driven, multicenter, international study of ERCP procedures conducted in both high-volume (> 1000 ERCPs/year) and low-volume (< 1000 ERCPs/year) centers in Southeastern Europe (Romania, Croatia, Italy, and Serbia). Invitations to participate in the study were issued to several institutions across Southeastern Europe that had collaborated previously on research projects in ERCP training, with the aim of canvassing a wide variety of endoscopic practices, including differences in training methods for novice endoscopists.

Endoscopists performing ERCPs at the participating centers were invited to prospectively document patient and procedure-related data using a standard report form (see **Appendix 1s** in the online-only supplementary material), which was designed to capture the technical aspects of the procedure as well as patient-related outcomes, including procedure-related adverse events and their outcomes. All consecutive ERCP procedures were documented, irrespective of trainee involvement; procedures without any trainee involvement served as a control group for the study.

The main study outcomes were technical success of the procedure and patient safety as assessed by the incidence of overall procedure-related adverse events.

Definitions

For the purpose of the study, trainees were defined as endoscopists working under direct supervision and/or endoscopists with fewer than 200 ERCP procedures performed independently (whichever applied). Technical success of the procedure was defined as the ability to achieve the planned diagnostic and/or therapeutic procedure in each individual case (e. g. correct positioning of stents, complete clearance of bile duct stones) and was assessed by the attending endoscopist at the end of each procedure. Technical difficulty of the procedure was graded according to the Schutz scale (▶ **Table 1**) [7]. Procedure-related adverse events were defined as any procedure-related complications that prolonged hospital stay and/or required additional medical or surgical interventions (e. g. surgery, additional endoscopic interventions, admission to the intensive care unit), and their severity was assessed using the Cotton scale, as previously described [8, 9]. Notably, although not usually considered as an adverse event per se, we included technical failure of the procedure resulting in prolonged hospital stay for the purpose of an endoscopic reintervention or a different therapeutic procedure (i. e. surgical or radiological drainage) as a procedure-related adverse event, to allow a more accurate assessment of the actual impact of trainee involvement.

▶ **Table 1** Modified Schutz scale for grading procedure difficulty [7].

Grade	Diagnostic	Therapeutic
1	CBD cannulation, brushing	Sphincterotomy, stones < 10 mm, stents for leaks and extrahepatic malignant strictures
2	Billroth II diagnostic, minor papilla cannulation	Stones > 10 mm, hilar stenting, benign biliary strictures
3	Manometry, Whipple, Roux-en-Y, intraductal endoscopy	Billroth II therapeutics, intrahepatic stones, pancreatic therapy

CBD, common bile duct.

Endoscopic procedures

ERCP was performed according to the standard procedure in each participating center, with respect to both the endoscopic procedure itself and the periprocedural medical care and surveillance. All centers used a wire-first approach with a triple-lumen sphincterotome for initial attempts at selective cannulation of the desired duct, with the notable exception of one center, which used contrast-guided cannulation with a double-lumen sphincterotome as the standard initial approach for selective cannulation.

Training protocol

Trainees worked under the direct supervision of an experienced endoscopist who supplied verbal as well as hands-on assistance and who could take over the procedure when deemed necessary. Accordingly, the degree of involvement of any trainee in any given procedure was entirely at the discretion of the supervisor, who acted on a case-by-case basis, according to best judgment and in the interest of ensuring patient safety throughout the procedure. Trainee involvement and technical success were divided and reported in the following categories, as detailed in the study form (**Appendix 1s**): 1) failed cannulation attempts; 2) cannulation of the desired duct; 3) partially

▶ **Table 2** Total number of procedures and percentage of trainee involvement per center.

Center (country)	Total no. of procedures, n	Procedures with trainee involvement, n (%)	Total number of endoscopists (experts/trainees), n
Colentina Hospital (Romania)	811	384 (47.3)	10 (4/6)
Policlinico Gemelli (Italy)	356	191 (53.7)	14 (5/9)
UHBC Zagreb (Croatia)	252	21 (8.3)	5 (4/1)
Cantacuzino Hospital (Romania)	201	201 (100)	2 (1/1)
Zadar Hospital (Croatia)	76	22 (28.9)	2 (1/1)
Belgrade University Hospital (Serbia)	147	4 (2.7)	2 (1/1)

successful procedure (including cannulation), but required hands-on assistance from the supervisor; 4) completed the procedure without hands-on assistance from the supervisor.

Data collection

Data were collected prospectively, using standard report forms completed by each supervising endoscopist at the end of the procedure. Follow-up of each patient was done by the attending endoscopist, with the aim of capturing procedure-related adverse events and patient outcome up to 30 days after the procedure. Although patients were subject to the individual post-procedure diagnostic and treatment protocols that were in place at each participating institution at the time of the study, as a general rule all patients undergoing ERCP were admitted and monitored for a minimum of 24 hours (mainly for reasons related to healthcare system reimbursement policies). Investigators were instructed to follow up on all patients for 30 days, using either patient chart review, telephone interviews, or both, in accordance with local protocols, regulations, and available infrastructure.

Statistical analysis

Data from all participating centers were collected in a central database and analyzed using SPSS v.16 (SPSS Inc. Chicago, Illinois, USA).

Data were analyzed in a two-step fashion. Bivariable analysis using the appropriate tests (chi-squared test, Student *t* test where continuous variables had a normal distribution, and Mann-Whitney *U* test where continuous variables had a non-normal distribution) was first carried out to identify potential risk factors for procedure-related adverse events and technical failure of the procedure. Multivariable analysis with logistic regression (using the enter method) was then carried out for the main study outcomes, including all variables that had a *P* level of <0.2 in the bivariable analysis. Based on clinical judgment and previously reported data, we also analyzed the potential interaction between procedure-related parameters in the model. Briefly, it was expected that trainee involvement would result in longer cannulation times (because the operator is inexperienced and requires more time to achieve deep duct cannulation), which, in turn, would mean that more procedures would be classified as difficult cannulation in the training group compared with the control group, as cannulation time >5 minutes is one of the defining criteria for difficult cannulation. Furthermore, prolonged cannulation time has been linked to more use of precut sphincterotomy. With regard to papilla anatomy, native papilla was expected to correlate with longer cannulation times, difficult cannulation, and use of precut sphincterotomy, as pre-existing sphincterotomy allows easy access to the desired ducts. As a result, we included all of these potential interactions in the multivariable models.

Sample size

The sample size was calculated for a 5% margin of equivalence in adverse events in the trainee group compared with the control group, from an estimated 10% overall, under the general assumptions of an equivalence trial (780×2 arms=1560 procedures, Sealed Envelope Ltd. 2012. Power calculator for binary outcome equivalence trial. Available from: https://www.sealedenvelope.com/power/binary-equivalence/ Accessed 8 September 2019). This also allowed for evaluation of a 5% margin of equivalence in cannulation success (using a standard cannulation success estimated at 95% in the control group) at a 0.05 significance level, with a beta of 0.9.

Ethical considerations

The study protocol was approved by the local ethics committee at each participating center, in accordance with the local and national regulations as well as the Declaration of Helsinki.

Results

A total of 16 independent operators and 21 trainees working at the 6 participating centers reported on 1843 ERCPs performed between October 2016 and October 2018. Trainees were involved in 822 procedures (44.6%), including 565 native papilla cases. Only 4 of the 21 trainees (19.0%) had been involved in >100 hands-on procedures prior to study inception, with the remaining trainees having limited or no experience in ERCP. The distribution of cases per center, as well the percentage of procedures with trainee involvement in each center, are illustrated in ▶ **Table 2**.

The most common indication for ERCP was the presence of common bile duct (CBD) stones (46.8%), followed by malignant strictures of the biliary tract (30.5%). Most patients were male (51.3%) and mean age in the patient population was 66.8 years

▶ **Table 3** General data about the patient population and the indication for endoscopic retrograde cholangiopancreatography (n = 1843).

Sex, male/female, n (%)	**946 (51.3)/897 (48.7)**
Age, mean (SD), years	66.8 (14.6)
Indication for the procedure, n (%)	
▪ CBD stones	863 (46.8)
▪ Malignant stricture of the bile duct	563 (30.5)
▪ Benign stricture of the bile duct	99 (5.4)
▪ Bile leak or trauma	57 (3.1)
▪ Stent exchange (for any indication)	168 (9.1)
▪ Chronic pancreatitis	40 (2.2)
▪ Other	41 (2.2)
Procedure difficulty (modified Schutz scale), n (%)	
▪ 1	1441 (78.2)
▪ 2	337 (18.3)
▪ 3	65 (3.5)
Type of sedation used, n (%)	
▪ No sedation	16 (0.9)
▪ Superficial sedation (midazolam)	136 (7.4)
▪ Deep sedation or general anesthesia	1691 (91.8)

SD, standard deviation; CBD, common bile duct.

(standard deviation 14.6). Native papilla cases were evenly distributed between the trainee and control groups (565 vs. 754; $P = 0.07$ chi-squared); however, there were significantly more grade 2 and 3 procedures in the control group than in the trainee group (25 % vs. 17.7 %; $P = 0.002$). General data about the patient population are further detailed in ▶ **Table 3**.

Trainee involvement and patient safety

In 270 cases (14.7 %) there was at least one procedure-related adverse event reported by the attending endoscopist. The most frequent adverse event reported was the technical failure of the procedure (89 cases, 4.8 %), followed by post-ERCP pancreatitis (PEP; 52 cases, 2.8 %) and cholangitis (45 cases, 2.4 %) (▶ **Fig. 1**). Most adverse events were either mild or moderate, but there were 19 severe adverse events (1.0 %) and 5 deaths occurring in the 30-day follow-up period (0.3 %); all of the fatalities occurred in patients with severe co-morbidities (American Society of Anesthesiologists score of 3 or more).

The overall incidence of adverse events was similar between the trainee group and the control group (14.7 vs. 14.6 %; $P > 0.99$ chi-squared) (▶ **Table 4**), with a difference of 0.1 percentage points between study groups (95 % confidence interval [CI] – 3.12 % to 3.38 %, within the 5 % predefined equivalence margin). However, there were significantly more moderate or severe adverse events in the control group compared with the trainee group (63 [6.2 %] vs. 28 [3.4 %]; $P = 0.01$ chi-squared).

▶ **Table 4** Comparison of patient and procedure-related characteristics between the two study groups.

Parameter	**Trainee group (n = 822)**	**Control group (n = 1021)**	***P* value**
Sex, male/female, n	398/424	548/473	0.04
Age, mean (SD), years	66.8 (14.1)	66.7 (14.9)	0.30
Technical success, %	760 (92.4)	957 (93.7)	0.30
Procedure difficulty, n (%)			
▪ Grade 1	676 (82.2)	765 (74.9)	0.002
▪ Grade 2	122 (14.8)	215 (21.1)	
▪ Grade 3	24 (2.9)	41 (4.0)	
Adverse events, n (%)			
▪ Any	121 (14.7)	149 (14.6)	> 0.99
▪ Mild	93 (11.3)	86 (8.4)	0.01
▪ Moderately severe	18 (2.2)	49 (4.8)	
▪ Severe, including fatalities	10 (1.2)	14 (1.4)	
Failure of the initial procedure and/or early reintervention required, n (%)	61 (7.4)	72 (7.1)	0.09
PEP, n (%)	25 (3)	27 (2.6)	0.58
Cholangitis, n (%)	16 (1.9)	29 (2.8)	0.29
Bleeding, n (%)	6 (0.7)	14 (1.4)	0.26
Perforation, n (%)	3 (0.4)	6 (0.6)	0.74
Other, n (%)	11 (1.3)	11 (1.1)	0.26

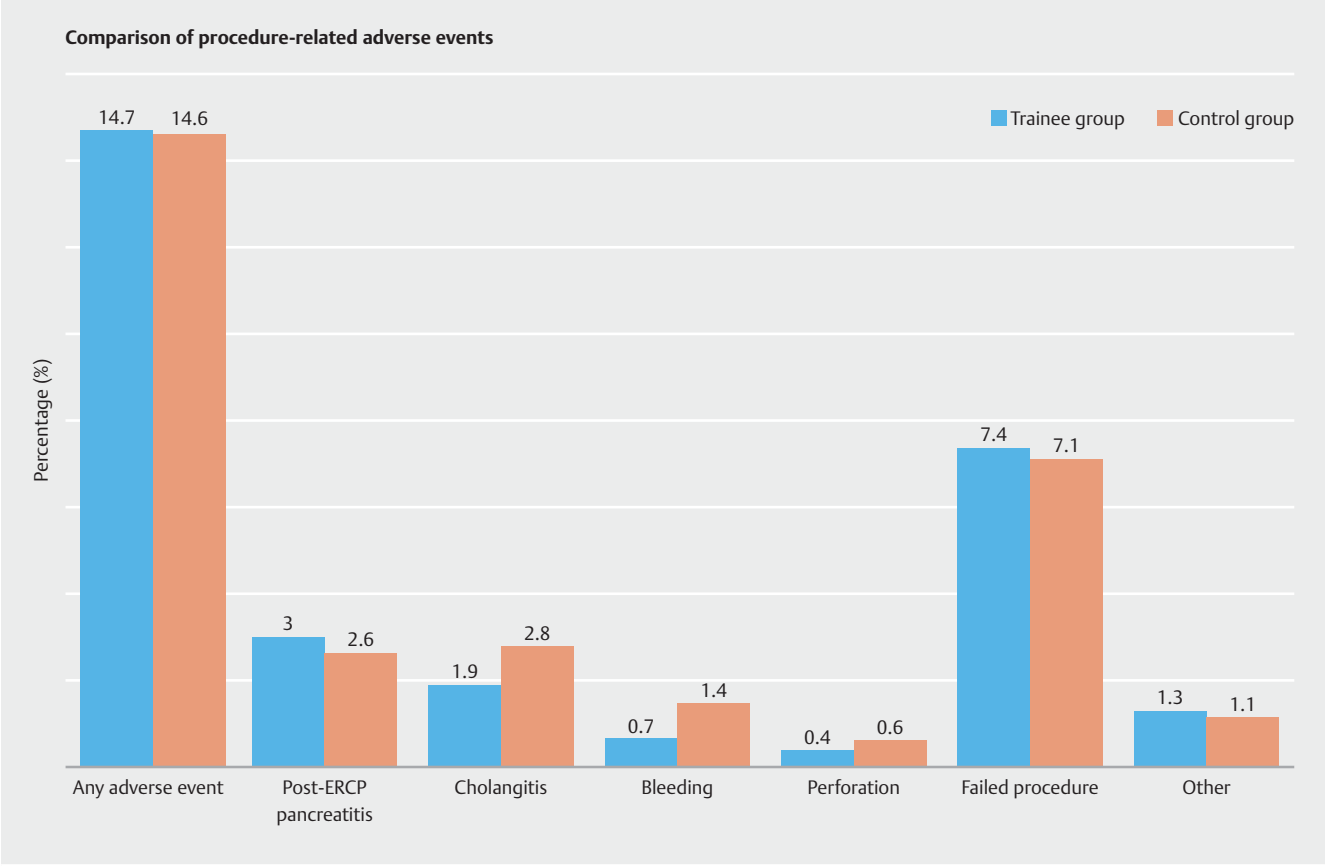
SD, standard deviation; PEP, post-endoscopic retrograde cholangiopancreatography.

Using bivariable analysis, bilirubin levels, presence of a native papilla, use of precut, difficult cannulation, procedure difficulty, indication for ERCP, time to cannulation, and participating center were selected as potential risk factors for the occurrence of a procedure-related adverse event (▶ **Table 5**) and included in the multivariable analysis model using logistic regression. We also adjusted for interaction between the relevant variables such as time to cannulation, difficult cannulation, native papilla, trainee involvement, and use of precut to exclude potential confounders in the model.

Finally, increased bilirubin levels, time to cannulation, and procedure difficulty level were shown to independently increase the risk of any procedure-related adverse event (▶ **Table 6**).

Trainee involvement and technical success of the procedure

Trainees were involved in 822 procedures and managed to successfully complete their respective procedures in 480 cases (58.4 %). In 150 cases (18.2 %), trainees required hands-on assistance from their supervisor to complete their procedure,



▶ **Fig. 1** Comparison of procedure-related adverse events across study groups.

and in 28 cases (3.4%) they could only selectively cannulate the desired duct, with a supervisor completing the rest of the procedure. In a further 164 cases (20.0%), trainees were unable to access the targeted duct, requiring an expert to take over to gain ductal access and, subsequently, perform the diagnostic or therapeutic procedure required in each case.

There was no difference in the incidence of failed cannulations between the trainee group (6.0%) and the control group (4.6%; $P=0.21$), including also subgroup analysis in native papilla cases only (7.7% vs. 5.6%; $P=0.14$).

The incidence of technical failure of the procedure, as appreciated by the attending endoscopist, was also similar between the trainee group and the control group (7.6% vs. 6.3%; $P=0.31$), with a calculated difference between the two groups of 1.3 percentage points (95%CI – 1% to 3.7%, falling within the 5% predefined equivalence margin).

On bivariable analysis (▶ **Table 5**), bilirubin levels, difficult cannulation, difficulty level of the procedure, use of precut, presence of a native papilla, time to cannulation, indication for the procedure, and participating center were all associated with an increased risk of procedure failure.

However, using multivariable analysis with logistic regression, only precut use, difficulty level, and time to cannulation were identified as risk factors for the technical failure of the procedure, after adjusting for ERCP indication and endoscopy center (▶ **Table 7**).

Discussion

The main finding of our study is that, although there is still a significant variation in training methods and, indeed, the practice of ERCP between endoscopy units, the differences in terms of procedure-related adverse events or technical failure between procedures with trainee involvement and those without trainee involvement do not seem clinically significant. In our trial, which was designed as an equivalence study, differences in the main outcome variables between the two groups did not exceed the predefined 5% margin.

Furthermore, although we noted a significant difference in the performance of individual endoscopy centers, the overall safety data across all participating centers was within the recommended targets [2] (PEP 2.8%, cholangitis 2.4%, post-sphincterotomy bleeding 1.1%, and perforations 0.5%), with a rate of severe adverse events of 1.0% and a 30-day mortality of 0.3%. Also in accordance with proposed European Society of Gastrointestinal Endoscopy guidelines, CBD cannulation rates were above the proposed 90% threshold, across both study groups, with overall technical success of the procedure also above 90%. Notably, this included a large number of procedures that were completed by trainees without any hands-on assistance from their supervisors (480 procedures out of a total of 822 procedures with trainee involvement). Our study was not powered to evaluate a potential increase in specific adverse

▶ **Table 5** Bivariable analysis of potential risk factors for adverse events and technical failure.

Potential risk factors	Any adverse event	No adverse event	*P* value	Technical success	Technical failure	*P* value
Sex, % female	49.4	48.3	0.74	48.6	45.6	0.52
Age, mean (SD), years	67.4 (13.9)	66.7 (14.7)	0.40	66.7 (14.6)	67.5 (13.6)	0.57
Native papilla, n (%)	216 (80.2)	1089 (69.4)	<0.001[1]	1195 (70.2)	104 (83.2)	0.001[1]
Use of precut, n (%)	93 (34.5)	183 (11.6)	<0.001[1]	223 (13.1)	53 (42.4)	<0.001[1]
Time to cannulation (<5/5–10/>10 min), n	93/38/136	1094/257/205	<0.001[1]	1164/285/240	16/7/101	<0.001[1]
Indication for ERCP, n (%)						
▪ CBD stones	91 (10.6)	770 (89.4)	<0.001[1]	823 (96.3)	31 (3.7)	<0.001[1]
▪ Benign stricture	14 (14.1)	85 (84.9)		88 (88.9)	11 (11.1)	
▪ Malignant stricture	124 (22.1)	438 (77.9)		491 (87.8)	69 (12.2)	
▪ Sclerosing cholangitis	1 (9.1)	10 (90.9)		11 (100)	0 (0)	
▪ Bile leak or trauma	12 (21.1)	45 (78.9)		50 (87.8)	7 (12.2)	
▪ Stent exchange	13(7.7)	155 (92.3)		164 (98.2)	3 (1.8)	
▪ Chronic pancreatitis	6 (15)	34 (85)		36 (92.3)	3 (7.7)	
▪ Other	8 (20)	32 (80)		39 (97.5)	1 (2.5)	
Procedure difficulty level (1/2/3), n	191/59/15	1216/277/50	0.03[1]	1316/307/52	83/29/11	0.001[1]
Trainee involvement, n (%)	120 (44.6)	701 (44.6)	>0.99	751 (44.1)	61 (48.8)	0.30
Participating center, n (%)						
▪ 1	150 (18.5)	660 (81.5)	<0.001[1]	737 (90.8)	74 (9.2)	<0.001[1]
▪ 2	36 (18)	164 (82)		186 (93)	14 (7)	
▪ 3	38 (15.7)	214 (84.3)		233 (92.4)	19 (7.6)	
▪ 4	6 (7.9)	70 (92.1)		73 (96.1)	3 (3.9)	
▪ 5	25 (7)	331 (93)		339 (98.5)	4 (1.5)	
▪ 6	14 (9.6)	132 (90.4)		137 (93.2)	10 (6.8)	
Difficult cannulation, n (%)	152 (43.1)	364 (23.2)	<0.001[1]	417 (24.5)	96 (77.4)	<0.001[1]
Bilirubin, median (range), mg/dL	5.7 (0.1–27.5)	2.2 (0.1–93.6)	<0.001[2]	2.38 (0.1–93.6)	7.95 (0.1–36)	<0.001[2]

SD, standard deviation; ERCP, endoscopic retrograde cholangiopancreatography; CBD, common bile duct.
[1] Using chi-squared test.
[2] Using Mann–Whitney *U* test.

events such as PEP, bleeding or cholangitis caused by trainee involvement. However, the consistently low incidence of these complications, well within the established guideline targets in both study groups, leads us to believe that any such existing relationship would have limited clinical significance in a real-life setting.

Training in endoscopy, and particularly in high-risk procedures such as ERCP, has been the focus of numerous studies, most of which were aimed at defining what represents competence in ERCP and how to best assess endoscopy trainees during and after their training period. The findings from these studies have shifted the paradigm in endoscopy training from a milestone-based approach (i.e. fulfilling a minimum number of procedures before independent practice is allowed) to a competence-based approach, requiring trainees to meet certain performance measures before being deemed competent. For example, the new Joint Advisory Group on Gastrointestinal Endoscopy accreditation system requires a trainee to show complication rate (death, transfusion requiring hemorrhage or perforation) of <5%, satisfactory completion of intended therapeutic procedure of >80%, and more than 75 procedures performed in the previous 12 months) [10].

However, the question of how these thresholds can be met in real life remains open to debate. Moreover, there are limited data about the outcome of ERCP procedures with trainee involvement, both in terms of technical success and, even more im-

▶ **Table 6** Risk factors for procedure-related adverse events identified using multivariable analysis.

Risk factor	OR	95 %CI
Increased bilirubin levels	1.01[1]	1.00 to 1.03
Time to cannulation (1)[2]	0.76	0.32 to 1.77
Time to cannulation (2)[2]	5.17	2.14 to 12.40
Difficulty level (1)[3]	1.53	1.05 to 2.23
Difficulty level (2)[3]	1.98	1.01 to 3.91
Trainee involvement	1.18	0.72 to 1.93

OD, odds ration; CI, confidence interval.
Covariates for the model included: endoscopy center, indication for endoscopic retrograde cholangiopancreatography, difficulty level, time to cannulation, difficult cannulation, use of precut, bilirubin levels, papilla anatomy, trainee involvement, as well as the following interactions: time to cannulation and trainee involvement; time to cannulation and use of precut; difficult cannulation and time to cannulation; difficult cannulation and papilla anatomy;, papilla anatomy and use of precut.
[1] Risk increase for every 1 mg/dL increase in bilirubin levels.
[2] OR provided for increase in time to cannulation from <5 to 5 – 10 min (1) and from 5 – 10 min to > 10 min (2), respectively.
[3] OR provided for increase in difficulty level from 1 to 2 (1) and from 2 to 3 (2), respectively.

▶ **Table 7** Multivariable analysis of risk factors for procedure failure.

Risk factor	OR	95 %CI
Use of precut	12.2	1.56 to 95.12
Difficulty level (1)[1]	2.01	1.14 to 3.51
Difficulty level (2)[1]	3.29	1.32 to 8.19
Time to cannulation (1)[2]	2.35	0.36 to 15.14
Time to cannulation (2)[2]	48.40	11.19 to 209.36
Endoscopy center[3]	0.25	0.10 to 0.71
Indication for ERCP (1)[4]	4.0	1.56 to 10.26
Indication for ERCP (2)[4]	6.79	2.20 to 20.89
Trainee involvement	1.23	0.36 to 4.17

OR, odds ratio; CI, confidence interval; ERCP, endoscopic retrograde cholangiopancreatography.
Covariates for the model included: endoscopy center, indication for ERCP, difficulty level, time to cannulation, difficult cannulation, use of precut, bilirubin levels, papilla anatomy, trainee involvement, and the following interactions: time to cannulation and trainee involvement; time to cannulation and use of precut; difficult cannulation and time to cannulation; difficult cannulation and papilla anatomy; papilla anatomy and use of precut.
[1] OR provided for increase in difficulty level from 1 to 2 (1) and from 2 to 3 (2), respectively.
[2] OR provided for increase in time to cannulation from <5 to 5 – 10 min (1) and from 5 – 10 min to > 10 min (2), respectively.
[3] OR reported for center number 4, which showed significantly higher technical success rates, using center number 1 as the reference category.
[4] OR reported using common bile duct stones as the reference category for this variable; (1) – benign bile duct stricture; (2) – bile leak or trauma.

portantly, concerning patient safety. Our study is the first European multicenter trial to address these key issues prospectively, gathering data from both referral centers and smaller endoscopy units from four countries in Southern and Eastern Europe.

With regard to patient safety, we were particularly stringent in defining the main outcome measure of procedure-related adverse events to include the partial or complete technical failure of the procedure, leading to prolonged hospital stay and need for reintervention. On multivariable analysis, we found increased difficulty level of the procedure and increased bilirubin levels to be risk factors for adverse events. Increased bilirubin levels are probably a surrogate marker of increased difficulty reflecting high grade or complex malignant strictures, such as hilar cholangiocarcinoma. Interestingly, an increase in cannulation time above 10 minutes was also found to increase the risk of adverse events, although trainee involvement, a recognized factor for delayed access to the bile duct [11], did not significantly increase the risk of adverse events in the multivariable analysis model.

Technical success of the procedure was achieved in over 90 % of the procedures in our study; however, there was a statistically significant difference between the participating centers, with one center in particular reporting extremely high success rates compared with the others. This finding is concordant with reports from other real-life reports, showing the large variation in performance between different centers [12, 13]. In our study, this finding could be explained by the fact that the outlier center carried out the highest number of procedures per year and had the largest number of expert endoscopists among the centers involved, although this was not explicitly analyzed in the study.

The main limitation of our study is its observational nature, which carries an inherent risk of selection bias with regard to the cases where trainees were involved compared with the control group. Although there were no significant differences concerning patient demographics between the two study groups, significantly more grade 2 and 3 cases were found in the control group compared with the trainee group. In addition, the degree of trainee and supervisor involvement in each case was not standardized, which could also potentially constitute a source of bias. However, we believe that the large number of included cases (> 1800 ERCPs), reflecting a broad spectrum of endoscopic practice with 6 centers and over 30 endoscopists involved in the study, minimizes the risk of bias and ensures an accurate reflection of the real-life situation of trainee involvement in ERCP procedures.

Another potential source of bias in our study is represented by the inherent differences in patient management, including the assessment and reporting of early and delayed adverse events, across centers in different countries. As none of the participating centers has access to a national patient database, it is conceivable that some adverse events, particularly delayed complications, might have been under-reported in our study. However, taking into account the large catchment area for each center involved in the trial, as well as the fact that all patients were admitted for at least 24 hours after the procedure,

we believe the risk of missing significant clinical outcomes to be very low in this particular cohort.

To date, there have been limited data reporting on the involvement of trainees in ERCP procedures, with somewhat contradictory findings and based on relatively small, single-center studies. While Frost et al. [14] showed that trainee involvement does not negatively influence either successful cannulation or time to cannulation, other data suggest that trainee involvement might actually increase the use of precut in order to facilitate access to the CBD [11]. With respect to patient safety, although available data [11, 14, 15] suggest that patient safety is not compromised by trainee involvement, these studies are usually underpowered for the analysis of complication rates.

Our findings underscore the fact that cannulation time, use of precut, and case complexity as assessed by the Schutz scale remain the most important risk factors for procedure failure and procedure-related adverse events, irrespective of trainee involvement in the procedure. We believe that future studies should focus on how a tailored approach to training, such as selecting low-risk cases for trainees in their initial phase of training, might further mitigate the risk of procedure-related adverse events while maximizing the benefits of hands-on training. Based on the evidence from this study, we believe trainers could be advised that some high-risk patients (i. e. those with high levels of bilirubin or those with grade 2 or 3 difficulty according to the Schutz scale) might not be adequate for inclusion on the procedure lists of less experienced endoscopists. Standardizing the approach to training has already been advocated by Kwek et al. [16], who propose some clear-cut criteria for situations in which the expert should take over from the trainee and continue the procedure, such as time on the papilla or exceeding a certain number of inadvertent pancreatic duct cannulations, with the aim of limiting the risk of complications. This approach could reasonably be expanded to incorporate high-risk features of the patient, as suggested by our study.

In conclusion, we have found that trainee involvement in ERCP procedures is safe, both in terms of procedure outcome and patient wellbeing. Future studies should focus on exploring whether procedure and training outcomes can be improved further by applying selection algorithms based on individual risk factors, to ensure that high-risk procedures are correctly identified and treated accordingly.

Acknowledgments

The authors would like to thank Paul Bălănescu, MD, PhD, who assisted with study design and statistical analysis of the database.

Competing interests

The authors declare that they have no conflict of interest.

References

[1] Jorgensen J, Kubiliun N. ASGE Training Committee. et al. Endoscopic retrograde cholangiopancreatography (ERCP): core curriculum. Gastrointest Endosc 2016; 83: 279–289

[2] Domagk D, Oppong KW, Aabakken L et al. Performance measures for ERCP and endoscopic ultrasound: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. Endoscopy 2018; 50: 1116–1112

[3] Voiosu T, Bălănescu P, Voiosu A et al. Measuring trainee competence in performing endoscopic retrograde cholangiopancreatography: a systematic review of the literature. United European Gastroenterol J 2019; 7: 239–249

[4] Wani S, Keswani RN, Han S et al. Competence in endoscopic ultrasound and endoscopic retrograde cholangiopancreatography, from training through independent practice. Gastroenterology 2018; 155: 1483–1494.e7

[5] Cheng CL, Sherman S, Watkins JL et al. Risk factors for post-ERCP pancreatitis: a prospective multicenter study. Am J Gastroenterol 2006; 101: 139–147

[6] Ekkelenkamp VE, Koch AD, Rauws E et al. Competence development in ERCP: the learning curve of novice trainees. Endoscopy 2014; 46: 949–955

[7] Schutz SM, Abbott RM. Grading ERCPs by degree of difficulty: a new concept to produce more meaningful outcome data. Gastrointest Endosc 2000; 51: 535–539

[8] Cotton PB, Lehman G, Vennes J et al. Endoscopic sphincterotomy complications and their management: an attempt at consensus. Gastrointest Endosc 1997; 37: 383–393

[9] Cotton PB, Eisen GM, Aabakken L et al. A lexicon for endoscopic adverse events: report of an ASGE workshop. Gastrointest Endosc 2010; 71: 446–454

[10] Isaacs P. Endoscopic retrograde cholangiopancreatography training in the United Kingdom: a critical review. World J Gastrointest Endosc 2011; 3: 30–33

[11] Voiosu T, Voiosu A, Benguş A et al. Trainee involvement increases precut rates and delays access to the common bile duct without an increase in procedure-related adverse events: a brave new world of ERCP training? Rom J Intern Med 2018; 56: 55–61

[12] Cotton PB, Romagnuolo J, Faigel DO et al. The ERCP quality network: a pilot study of benchmarking practice and performance. Am J Med Qual 2013; 28: 256–260

[13] Kapral C, Duller C, Wewalka F et al. Case volume and outcome of endoscopic retrograde cholangiopancreatography: results of a nationwide Austrian benchmarking project. Endoscopy 2008; 40: 625–630

[14] Frost JW, Kurup A, Shetty S et al. Does the presence of a trainee compromise success of biliary cannulation at ERCP? Endosc Int Open 2017; 5: E559–E562

[15] Iida T, Kaneto H, Wagatsuma K et al. Can trainees safely perform endoscopic treatments for common bile duct stones? a single-center retrospective study Intern Med 2018; 57: 923–928

[16] Kwek BE, Ang TL, Teo EK et al. Making ERCP training safe: a protocol-based strategy to minimize complications during selective biliary cannulation. J Interv Gastroenterol 2012; 2: 66–69