Human Reproduction Open, pp. 1–15, 2020
doi:10.1093/hropen/hoaa026

human reproduction open

ESHRE PAGES

ESHRE Clinical Embryologist certification: the first 10 years†

Borut Kovačič [1,*], Fernando J. Prados[2], Catherine Plas[3], Bryan J. Woodward[4], Greta Verheyen[5], Liliana Ramos[6], Sirpa Mäkinen[7], Susanna Jamina Apter[8], Francesca Vidal[9], Søren Ziebe[10], M. Cristina Magli[11], Kersti Lundin[12], Arne Sunde[13], and Carlos E. Plancha[14]

[1]Department of Reproductive Medicine, University Medical Centre Maribor, Maribor, Slovenia [2]Independent researcher, Madrid, Spain [3]ESHRE Central Office, Grimbergen, Belgium [4]X&Y Fertility, Leicester, UK [5]University Hospital, Vrije Universiteit Brussel, Centre for Reproductive Medicine, Brussels, Belgium [6]Department of Obstetrics and Gynaecology, Radboud University Medical Centre, Nijmegen, The Netherlands [7]Ovumia Fertinova, Helsinki, Finland [8]Livio Fertilitetscentrum, Stockholm, Sweden [9]Unitat de Biologia Cellular, Universitat Autonoma de Barcelona, Barcelona, Spain [10]The Fertility Clinic, Section 4071, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark [11]SISMER Reproductive Medicine Unit, Bologna, Italy [12]Reproductive Medicine, Sahlgrenska University Hospital, Gothenburg, Sweden [13]Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway [14]Instituto de Histologia e Biologia do Desenvolvimento, Faculdade de Medicina, Universidade de Lisboa, Portugal and CEMEARE, Lisboa, Portugal

*Correspondence address. Department of Reproductive Medicine, University Medical Centre Maribor, Maribor, Slovenia.Tel: +386 2 321 2192; E-mail: borut.kov@ukc-mb.si https://orcid.org/0000-0003-3911-2408

Submitted on March 16, 2020; editorial decision on April 20, 2020

STUDY QUESTION: What has the ESHRE programme 'ESHRE Certification for Clinical Embryologists' achieved after 10 years?

SUMMARY ANSWER: The post-exam analysis showed a pass rate of 60% for Clinical and 50% for Senior Clinical Embryologists and a high level of internal consistency of all exams, leading to a total of 773 certified Clinical and 493 Senior Clinical Embryologists over the decade.

WHAT IS KNOWN ALREADY: In an ESHRE survey on the educational and professional status of Clinical Embryology in Europe, it was found that education of laboratory personnel working in the field of assisted reproduction is highly variable between countries. In 2008, ESHRE introduced a programme, curriculum and certification in the field of Clinical Embryology. Knowledge gained by postgraduate study of recommended literature, following a clear curriculum, is verified by a written two-level exam for obtaining a certificate for Clinical (basic) or Senior Clinical (advanced) Embryologists. With a total of 1266 certificates awarded over a period of 10 years and recognition by the Union Européenne des Médecins Spécialistes and their Council for European Specialists Medical Assessment, the ESHRE Clinical Embryology exams have become an internationally recognized educational standard in the field of Clinical Embryology.

STUDY DESIGN, SIZE, DURATION: A retrospective analysis of all applications for ESHRE Clinical (2009–2018) and Senior Clinical Embryologist Certification (2008–2018) and exam results of the first decade was carried out by the Steering Committee for Clinical Embryologist Certification.

PARTICIPANTS/MATERIALS, SETTING, METHODS: A total of 2894 applications for ESHRE Certification for Clinical Embryologists and the results of 10 exams for the Clinical (1478 candidates) and 11 exams for Senior Clinical (987 candidates) levels were analysed. A detailed post-exam retrospective analysis was performed regarding difficulty, discrimination and reliability levels of 1600 multiple-choice questions (MCQs) with a single best answer among four options, from eight different curriculum topics (Basic cell biology, Genetics, Developmental biology, Female reproduction, Male reproduction, IVF laboratory, Cryopreservation and Laboratory management), representing the core theoretical knowledge of Clinical Embryology. Difficulty levels of the MCQs were subsequently compared regarding each topic and each yearly exam. The participation and success rates in the ESHRE Clinical Embryology exams were also assessed in terms of the educational and geographic backgrounds of candidates.

MAIN RESULTS AND THE ROLE OF CHANCE: Over the 10 years studied, the mean pass rate for the Clinical Embryologist exam was 60% (range 41–86%), and for the Senior Clinical Embryologist exam was 50% (range 34–81%). On average, 63% European candidates

†ESHRE pages content is not externally peer reviewed. The manuscript has been approved by the Executive Committee of ESHRE.

and 35% non-European candidates passed the Clinical Embryologist exam, while 52% European candidates and 31% non-European candidates passed the Senior Clinical Embryologist exam. The candidates' educational level impacted on the success of the Clinical Embryologist exam but not of the Senior Clinical Embryologist exam. The mean difficulty indices by study topic showed that in the period of 10 years, there were no statistically significant differences between topics, for either the Clinical or Senior Clinical Embryologist exams. However, the overall exam difficulty varied between years. Reassuringly, the exam MCQ discrimination and reliability indices always showed a high level of internal consistency in all exams.

LIMITATIONS, REASONS FOR CAUTION: Some data from the initial ESHRE certification programme were not obtained electronically, in particular data for education, implying tables and figures reflect the specified valid data periods. Several countries exhibit different study profiles for those working in ART laboratories, such that laboratory technicians/technologists predominate in some countries, while in others only biologists and medical doctors are allowed to work with human embryos. Such differences could consequently affect the exam performance of candidates from specific countries.

WIDER IMPLICATIONS OF THE FINDINGS: The ESHRE exams on Clinical Embryology are the most widely, internationally accepted tests of knowledge in the rapidly growing area of human reproduction. Clinical Embryology is increasingly recognized as a specific discipline for scientific staff who are collaborating closely with clinicians in managing human infertility through medically assisted reproduction. The analysis of the first 10 years of application of a two-level exam for Clinical Embryology shows a consistent high quality and reliability of the exam and MCQs used. These results represent an important follow-up of the quality of the ESHRE Certification programme for Clinical Embryologists, and convincingly position Clinical Embryology in the wider group of health disciplines that are harmonized through professional bodies such as ESHRE and European Board & College of Obstetrics and Gynaecology. The exams provide a clear step towards the increasing professional recognition and establishment of Clinical Embryology within health systems at both European and international level.

STUDY FUNDING/COMPETING INTEREST(S): No competing interest. All costs of the Steering Committee meetings were covered by ESHRE.

Key words: clinical embryology / education / embryologist certification / exam / post-exam item analysis / ESHRE

Introduction

From the late 1970s/early 1980s, professionals experienced in gamete/embryo culture or cell culture systems co-operated with clinicians to set up the first clinical ART laboratories. These pioneers included scientists from basic biological or biomedical fields and veterinarians with practical experience in animal reproduction, as well as individuals with several other backgrounds. Reproductive medicine has since experienced an exponential development and growth, particularly due to improvements in the laboratory aspects of medically assisted reproduction (MAR).

It soon became evident that the quality of laboratory work was critical to the success of MAR treatment. The need for specific knowledge and skills for manipulating human gametes and embryos, performing high-technology procedures and consistently ensuring high quality culture conditions, led to the establishment of a new discipline in laboratory medicine: Clinical Embryology.

The need for wider recognition of the field of Reproductive Medicine encouraged ESHRE together with the European Board & College of Obstetrics and Gynaecology (EBCOG) to develop an educational system enabling medical clinicians to be granted the title of subspecialists in Reproductive Medicine (Calhaz-Jorge *et al.*, 2015). For Clinical Embryologists working in MAR centres; however, the situation was more complex, particularly given the high heterogeneity of academic levels (BSc, MSc or PhD) and scientific backgrounds of these professionals working in healthcare (Alpha Scientists in Reproductive Medicine, 2015; Kovačič *et al.*, 2015).

Professional bodies in some European countries have provided national recommendations regarding qualifications for Clinical Embryologists (Hughes and Association of clinical embryologists, 2012; Trávník *et al.*, 2013); however, in most countries, no formal recognition of this profession exists. In order to bring further clarification and harmonization, ESHRE decided to establish a European certification/qualification programme with a comprehensive scientific curriculum. In fact, the internationally recognized technical and scientific competences of ESHRE positioned the society to develop such a system, aimed at providing wide and formal recognition of the status of Clinical Embryologists. In 2006, it was decided to create an ESHRE Task Force for the Certification of Clinical Embryologists.

The Certification Task Force established contact with several professional and scientific bodies in order to seek input in the curriculum topics. ESHRE consulted Alpha Scientists in Reproductive Medicine (ALPHA), Asociación para el Estudio de la Biología de la Reproducción (ASEBIR, Spain), Arbeitsgemeinschaft Reproduktionsbiologie des Menschen (AGRBM, Germany), Association of Clinical Embryologists (ACE, UK) and Nordic IVF Laboratory Society (NILS), Vereniging voor Klinisch Embryologen (KLEM, The Netherlands), Vlaamse Vereniging voor Klinische Embryologie (VVKE, Belgium), Association Belge des Embryologistes Francophones (ABEF, Belgium), Panhellenic Association of Clinical Embryologists (PEKE, Greece), Embryologenforum Austria (EFA) and key embryologists from Portugal. This consultation resulted in the establishment of the ESHRE Embryology Certification Committee (EmCC) who developed a two-level self-educational certification programme for Clinical Embryologists (basic level) and Senior Clinical Embryologists (advanced level).

The proposed application system included a background curriculum, evidence of academic levels, a logbook showing a minimum number of performed key procedures in the ART laboratory, and a formal exam to test the candidates' theoretical knowledge. The major objective

was to test the Clinical Embryologists' understanding of the fundamental biology underpinning the technical and clinical aspects of MAR.

The ESHRE Certification for Clinical Embryologists programme started in 2008 with a fast-track certification for Senior Clinical Embryologists having served for at least 10 years, being head of an ART laboratory and holding a PhD in a relevant discipline. Following this, the first ESHRE exam was held at the ESHRE Annual Meeting in Barcelona 2008, where only candidates for the Senior Clinical Embryologist exam were allowed to participate. From 2009 onwards, the system was also open to Clinical Embryologists.

In the period 2008–2018, the certification programme underwent several improvements by introducing electronic exam evaluation (2011), external validation of the academic diploma (2012) and an online application platform (2015). Certification was also opened for non-European embryologists in 2012. Furthermore, in 2015, ESHRE introduced an online system for monitoring Continuing Professional Development (CPD) for all certified embryologists. In a pilot project, one long-distance on-site exam for a small group of 22 candidates was organized in 2018 in India, simultaneously with the main exam that was held in Geneva. In 2016, ESHRE applied for validation of the entire certification process by the Union Européenne des Médecins Spécialistes (UEMS) and their Council for European Specialists Medical Assessment (CESMA). As a result, UEMS-CESMA monitored the ESHRE Certification for Clinical Embryologists exams in Geneva in 2017 and the ESHRE exams for embryologists became officially recognized in accordance with the UEMS standards.

The current report presents a detailed analysis of the first decade of certification of Clinical Embryologists including the number of candidates, their educational background and the pass rate for the Clinical or Senior Clinical Embryologist level exams. Data regarding the difficulty, discrimination and reliability levels of exams and questions from the various curriculum topics of Clinical Embryology are also provided. In addition, the variation in participation and success rate between different educational and geographic backgrounds is presented. This article aims to pave the way for further improvements in the ESHRE Certification for Clinical Embryologists programme.

Material and methods

Steering committee and exam questions

A team of experts selected by ESHRE formed the EmCC, with responsibility for the self-educational Clinical Embryology programme. The EmCC presented a detailed curriculum, recommended appropriate teaching materials, prepared exam questions, regularly opened a call for exam participation, verified the candidates' eligibility, organized exams, evaluated the exam results and addressed complaints. The EmCC comprised eight members, representing different areas of Europe and speciality topics, and one ESHRE administrative officer.

The basis of the self-educational programme was a curriculum that identified eight main topics from Clinical Embryology: Basic cell biology, Genetics, Developmental biology, Female reproduction, Male reproduction, IVF laboratory, Cryopreservation and Laboratory management.

Appropriate updated/revised literature was recommended to candidates preparing for the exam. Exam questions were based on the recommended literature and good laboratory and clinical practice.

Multiple-choice questions (MCQ) of the single best answer (SBA) type were developed for the exam, reviewed and approved by all EmCC members, and requiring one correct answer to be selected among the four answer options. All MCQs were divided into three groups according to the estimated level of knowledge required: specific for Clinical Embryologists, specific for Senior Clinical Embryologists and common to for both (Clinical and Senior Clinical Embryologists). The MCQs have been developed and updated over time, based on evolving knowledge and post-exam performance analysis.

All questions were secured electronically, each of them containing information about its topic, predicted difficulty and unique serial number. Once used, each question was linked also with data about the year of use, percentages of chosen answers and difficulty index (see below) based on post-exam analysis.

Curriculum and logbook

The Clinical Embryologist exam was set for laboratory staff who fulfilled the following requirements at the date of the application deadline: at least a BSc degree in natural/life sciences, at least 3 years hands-on experience with human gametes and embryos in an ART laboratory, and a minimum of 50 hands-on procedures per laboratory topic, as specified below. The Senior Clinical Embryologist exam was set for candidates with either a MSc or PhD, at least 6 years hands-on experience in a ART laboratory, and a minimum of 50 hands-on procedures per laboratory topic, as specified below.

Applicants were informed about the curriculum, requirements, application procedure and type of exam via dedicated webpages added to the ESHRE website. Applications were open for 2 months from mid-October until mid-December. From 2015 onward, an online application platform was developed and the applicants' highest university degree was verified by an independent European organization: the Netherlands Universities Foundation For International Cooperation (NUFFIC). The validation process was co-ordinated via ESHRE. Diplomas that did not fit the criteria of an official national university degree were not accepted.

A logbook was required to provide a record of the number of personally performed procedures in nine laboratory tasks: oocyte retrieval, semen analysis, semen preparation, IVF insemination, ICSI, zygote and embryo evaluation, embryo transfer, cryopreservation of oocytes/embryos, thawing/warming of oocytes/embryos. A minimum number of 50 cases per topic was required in a period of 3 years for Clinical Embryologist applicants or 6 years for Senior Clinical Embryologist applicants (see ESHRE link for further specifications).

The information from the application and logbook(s) had to be verified and signed by the supervisor(s) from the respective MAR centres. Two references were required to support the application, preferably being ESHRE members and/or ESHRE Certified Senior Clinical Embryologists. If the candidate had been working at several clinics, separate signed logbooks were required.

Exam and marking

The exams were held at every ESHRE Annual Meeting. Each exam comprised 100 MCQs of the SBA type. Candidates received 50 level-specific questions and 50 questions common to both levels. The official exam language was English; however, for the first three exams only, the exam was also provided in the languages of the host country of

the ESHRE Annual Meeting (Spanish, Dutch and Italian). The exam time was 2 h and 15 min. The applicants were notified in advance of the exam that they needed to pass a threshold of 66% of the answers correctly to pass the exam (see Discussion section).

Marking was performed electronically and verified manually by ESHRE officers to ensure accuracy. The marking software provided information on candidate performance for each question. The EmCC performed a post-exam analysis (see below). Questions with a high proportion of incorrect answers were reconsidered and, if appropriate, accepted with more than one correct answer, or removed. A final check was always made prior to informing candidates of their performance.

Study design and outcome measurements

All applications and exam results for ESHRE Clinical (2009–2018) and Senior Clinical Embryologist certification (2008–2018) were analysed retrospectively. The demographic data of candidates were collected from their application files. From 2015 onwards, the application platform was updated to collect more detailed information on study fields and academic degrees. The online application system simplified confirmation of the eligibility of applicants. A maximum of 300 candidates was permitted to apply per exam.

Information about exam performance was taken from the EmCC database, in which relevant data (number of applications, valid applications, pass rate, mean score, median score and mean difficulty level) were collected annually for quality control.

The aim of this retrospective study was to present applicants' educational background and the pass rates according to participant group (new applicants vs resits, Europeans vs non-Europeans and academic degrees). From post-exam analysis, the difficulty, discrimination and reliability indices were calculated. Answer options that should not be marked by the candidate are defined as distractors. A distractor analysis was also performed (see below). Additional outcome measure parameters included difficulty levels of specific topics and number of successful/failed exam attempts according to the registered country of candidates. All study outcomes are presented separately for Clinical and Senior Clinical Embryologist exams.

Statistical analysis

For exams from 2009 to 2010, the evaluation and the analysis were performed manually. For exams from 2011 onwards, the analysis was performed electronically using the relevant computer software (Remark Office OMR 2014, Gravic, Inc. Malvern, PA, USA). The IBM SPSS Statistics (version 22) software package was used for χ^2 test and analysis of variance (ANOVA). Post-exam analysis of 1600 MCQs from both types of exams was based on 'item analysis' indices. This provided a post-exam p-statistic (difficulty index) (Downing *et al.*, 1995), a generalized upper-lower item discrimination index (Richardson and Stalnaker, 1933; Ebel, 1954) for each question, and general reliability of assessment (also known as internal consistency reliability) presented as Kuder–Richardson (KR-20) parameter (Kuder and Richardson, 1937) (see below).

The p-statistic was a proxy for the difficulty of each question. Each p-value was calculated as the percentage of candidates providing a correct response to a question. The percentage was then divided by 100 and converted to a 'difficulty index'. Using the classical terminology in the field, the difficulty index therefore results from the proportion of correctly answered MCQs. It therefore reflects the 'easiness' of the MCQ: an MCQ with a higher difficulty index value means an easier question. Ideally, MCQs should have p-values (difficulty indices) that range between 0.2 (20% correct answers) and 0.8 (80% correct answers) (Dixon, 1994). One-way ANOVA allowed for comparison of the mean difficulty indices of different topics and between exams.

Questions from each curriculum topic were categorized into 10 groups according to difficulty, from very easy (level 1, difficulty index range 0.9–1) to very difficult (level 10, difficulty index range 0–0.1). In detail, the difficulty groups were: 1 (0.9–1), 2 (0.81–0.9), 3 (0.71–0.8), 4 (0.61–0.7), 5 (0.51–0.6), 6 (0.41–0.5), 7 (0.31–0.4), 8 (0.21–0.3), 9 (0.11–0.2) and 10 (0–0.1).

The post-exam discrimination analysis was computed from equal-sized upper and lower scoring candidates (27% highest and lowest scoring candidates) according to Kelley (1939). Values above 0.2 were regarded as optimal (Ebel, 1954).

KR-20 is the variance of the total scores of all the candidates taking the test. KR-20 values range from 0 to 1. A high KR-20 indicated reliable (consistent) student scores, with a KR-20 score of 0.60 or higher being desirable.

To compare pass rates between candidates of different educational background, the χ^2 test was used. A *P*-value of <0.05 was considered as statistically significant. Difference between countries in number of failed and successful exam attempts are presented graphically.

Results

Exam analysis: applications, pass rates and mean scores

The data for Fig. 1 were obtained from the online application platform for the last three exams (2016–2018). Prior to 2016, the educational background of candidates was verified from the submitted documentation, but this information was not analysed. The applicants from both exam levels had university degrees from different life science study fields (Fig. 1). The majority studied biology (43.5%), biotechnology (15.1%) or biomedical sciences (10.1%), accepting that different countries may use different nomenclature for the topic of a study field. The representation of study fields did not differ between applicants for Clinical and Senior Clinical Embryologist exams, with the exception of Laboratory Technology, which was more frequently represented in the Clinical Embryologist exam (6.1% vs 1.7%, respectively $P = 0.005$).

Data regarding applications for the Clinical Embryologist exam are presented in Table I. In the period from 2009 to 2018, a total of 1611 applications were received, of which 92% (1478 candidates) met the application criteria. Altogether, 1289 candidates attended the exam, with 89% participants from European and 11% from non-European countries. The mean pass rate for the Clinical Embryologist exam was 60%, ranging from 41% to 86%. On average, 80% of the candidates attended the exam for the first time and reached the mean pass rate of 63%. Twenty percent of applicants were resits, with the lower mean pass rate of 49%. On average, 63% European candidates and 35% non-European candidates passed the Clinical Embryologist exam. The mean score achieved by all participants in the Clinical Embryology exams was 68/100 and by those who passed exams was 76/100.

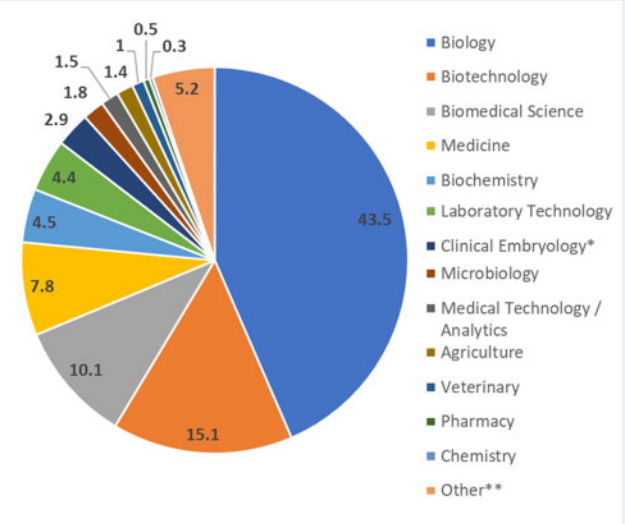

Figure 1 Frequency distribution (%) of examinees' basic study fields, 2016–2018. Clinical Embryology* (University Master; not titulo propio—a degree, offered by a single university, which is not officially recognized by the national educational system), other** (Bioanalytical Science, Bioengineering, Environmental Science, Forensic Science, Human Biology, Medical Informatics, Molecular Biology and Genetics, Physiology, Zoology, etc.).

The second part of Table I presents the data of the Senior Clinical Embryologist exams. In total, 1283 candidates applied, 90% (1158 applicants) were accepted and finally 987 applicants attended the exam. The mean pass rate was 50%, ranging from 34% to 81%. On average, 77% were first-attempt candidates and 23% were resit candidates (50% and 39% mean pass rate, respectively). Altogether, there were 89% European and 11% non-European participants (52% and 31% mean pass rate, respectively). The mean score achieved in Senior Clinical Embryology exams was 59/100 and in those who passed exams it was 73/100.

The frequency of scores achieved in all Clinical Embryology and Senior Clinical Embryology exams is presented graphically (Fig. 2a and b).

The candidates' educational level impacted on the success of the Clinical Embryologist exam (Table II). Candidates with a PhD obtained a higher pass rate (67%) than candidates with a MSc (55%) ($P = 0.022$) or BSc (48%) ($P < 0.001$). However, for the Senior Clinical Embryologist exam, there was no difference in pass rates among candidates with a PhD or MSc (42% and 48%, $P = 0.14$). In line with the expected higher difficulty of the Senior Clinical versus the Clinical Embryology exams, the detected difference in pass rates between PhD groups (67% vs 42%, $P < 0.001$) was expected. However, candidates with an MSc degree obtained similar pass rates in both types of exams (55% vs 48%).

The number of attempts (tries by the candidates, that can be successful or not) according to country, the exam level and performance is shown in Fig. 3a and b. The highest number of attempts for the Clinical Embryologist exam came from Poland ($n = 131$), Italy ($n = 120$), Spain ($n = 100$), Denmark ($n = 99$) and Sweden ($n = 81$); and for the Senior Clinical Embryologist exam came from Spain ($n = 251$), Italy ($n = 96$), Poland ($n = 69$), Germany ($n = 65$) and Greece ($n = 54$). The success rate per country did not correlate with the number of attempts.

Question analysis

From 2011 to 2018, the number of MCQs with optimal difficulty index (range from 0.2 to 0.8) increased from around 60% in the Clinical Embryologist exam and 70% in the Senior Clinical Embryologist exam to around 80% in both (Table III). The overall mean difficulty indices of Clinical Embryologist and Senior Clinical Embryologist exams were 0.68 ± 0.19 (range from 0.63 to 0.74 between years) and 0.64 ± 0.21 (range from 0.59 to 0.7), respectively (Table III).

The specific questions for Clinical Embryologists and Senior Clinical Embryologists showed mean difficulty indices of 0.71 and 0.56, respectively. The common questions used for both the Clinical Embryologist and Senior Clinical Embryologist exams demonstrated mean difficulty indices of 0.62 and 0.65, respectively (data not shown in tables).

Supplementary Table SI presents the mean difficulty indices of both types of exams, per years and per specific topic. The analysis of the mean difficulty indices per study topic showed that in the period of 10 years, there were no statistically significant differences between topics, for either the Clinical or Senior Clinical Embryologist exams. However, the overall exam difficulty varied between years ($P < 0.001$).

Most MCQs from Clinical Embryologist exams had difficulty levels between 2 and 4, while most Senior Clinical Embryologist exam questions were between 2 and 5 (Fig. 4a and b).

The majority of MCQs from the Clinical Embryologist exams (70%) and the Senior Clinical Embryologist exams (68%) demonstrated a sufficiently high discrimination index (≥ 0.2). The mean discrimination indices were 0.27 ± 0.12 (Clinical Embryologist exam) and 0.26 ± 0.14 (Senior Clinical Embryologist exam). The annual discrimination index never dropped below 0.2 (Table III).

The relation between the difficulty and discrimination indices is shown in scatterplots for 1600 MCQs from both exams (Fig. 5a and b). The blue area indicates the optimal discrimination and difficulty ranges. MCQs in the pink area are very difficult with low discrimination, whilst MCQs in the green area are very easy with low discrimination. Most MCQs are in the blue area. In the Senior Clinical Embryologist exams, many MCQs showed low discrimination, which is characteristic of making a random choice by the candidates in choosing the right answer. MCQs in the red area or having a negative discrimination index were set aside for future improvement. Only 1.8% of MCQs in the Clinical Embryologist exams and 3.8% in the Senior Clinical Embryologist exams had negative discrimination indices.

There were 4800 answer options that should not have been chosen by the candidate from 1600 MCQs (Tables III). These answer options are termed 'distractors'. If a MCQ is good quality, then all distractors should be selected by some candidates. If a MCQ is poorer quality, then one or more of the distractors should not be selected by all candidates. In such instances, poorer quality MCQs are less discriminative, as the answer options are 'low-quality' distractors, also known as 'non-distractors'. Thus, the quality of each distractor can modulate the difficulty of an MCQ. In both levels of exam, a trend for an increasing number of better quality distractors was observed from 2011 to 2018. Consequently, the number of MCQs with 'non-distractors' has

Table I Results of ESHRE exams.

Exams for Clinical Embryologists

Where exam held	Amsterdam 2009	Rome 2010	Stockholm 2011	Istanbul 2012	London 2013	Munich 2014	Lisbon 2015	Helsinki 2016	Geneva 2017	Barcelona 2018	Total
Applications											
Applications received	115	161	156	135	143	171	184	149	176	221	1611
Valid applications (%)	99 (86)	143 (89)	141 (90)	122 (90)	134 (94)	160 (94)	167 (91)	137 (92)	168 (95)	207 (94)	1478 (92)
Participants											
All participants	88	127	121	107	120	135	149	121	142	179	1289
New applicants (%)	88 (100)	116 (91)	95 (79)	88 (82)	79 (66)	103 (76)	96 (64)	99 (82)	117 (82)	149 (83)	1030 (80)
Resits (%)		11 (9)	26 (21)	19 (18)	41 (34)	32 (24)	53 (36)	22 (18)	25 (18)	30 (17)	259 (20)
European (%)	88 (100)	127 (100)	121 (100)	107 (100)	120 (100)	123 (91)	120 (81)	102 (84)	108 (76)	130 (73)	1146 (89)
Non-European (%)						12 (9)	29 (19)	19 (16)	34 (24)	49 (27)	143 (11)
Pass rate											
All participants (%)	76 (86)	97 (76)	87 (72)	55 (51)	77 (64)	65 (48)	108 (72)	50 (41)	70 (49)	88 (49)	773 (60)
New applicants (%)	76 (86)	87 (75)	71 (75)	46 (52)	51 (65)	57 (55)	69 (72)	49(49)	65 (46)	76 (51)	647 (63)
Resits (%)		10 (91)	16 (62)	9 (47)	26 (63)	8 (25)	39 (74)	1 (5)	5 (20)	12 (40)	126 (49)
European (%)	76 (86)	97 (76)	87 (72)	55 (51)	77 (64)	62 (50)	94 (78)	43 (42)	60 (56)	72 (55)	723 (63)
Non-European (%)						3 (25)	14 (48)	7 (37)	10 (29)	16 (33)	50 (35)
Scores											
Mean score (%)	76	73	71	66	69	65	72	63	65	65	68
Mean score of participants passed (%)	79	77	77	73	76	74	78	75	75	75	76

Exams for Senior Clinical Embryologists

Where exam held	Barcelona 2008	Amsterdam 2009	Rome 2010	Stockholm 2011	Istanbul 2012	London 2013	Munich 2014	Lisbon 2015	Helsinki 2016	Geneva 2017	Barcelona 2018	Total
Applications												
Applications received	170	140	101	72	91	120	110	116	102	111	150	1283
Valid applications (%)	152 (89)	115 (82)	89 (88)	62 (86)	81 (89)	111 (92)	107 (97)	101 (87)	93 (91)	105 (95)	142 (95)	1158 (90)
Participants												
All participants	145	100	74	51	66	94	91	80	78	85	123	987
New applicants (%)	145 (100)	89 (89)	49 (66)	30 (59)	44 (67)	53 (56)	65 (71)	48 (60)	63 (81)	72 (85)	103 (84)	761 (77)
Resits (%)		11 (11)	25 (34)	21 (41)	22 (33)	41 (44)	26 (29)	32 (40)	15 (19)	13 (15)	20 (16)	226 (23)
European (%)	145 (100)	100 (100)	74 (100)	51 (100)	61 (92)	75 (80)	75 (82)	73 (91)	62 (79)	91 (76)	101 (82)	882 (89)
Non-European (%)					5 (8)	19 (20)	16 (18)	7 (9)	16 (21)	20 (24)	22 (18)	105 (11)
Pass rate												
All participants (%)	118 (81)	40 (40)	37 (50)	19 (37)	26 (39)	43 (46)	31 (34)	58 (73)	28 (36)	44 (52)	49 (40)	493 (50)
New applicants (%)	118 (81)	38 (43)	29 (59)	12 (40)	24 (55)	22 (42)	19 (29)	17 (77)	27 (43)	35 (41)	43 (35)	384 (50)
Resits (%)		2 (18)	8 (32)	7 (33)	2 (9)	21 (51)	12 (46)	21 (66)	1 (7)	9 (69)	6 (30)	89 (39)
European (%)	118 (81)	40 (40)	37 (50)	19 (37)	24 (39)	36 (48)	26 (35)	53 (73)	24 (39)	35 (38)	48 (48)	460 (52)
Non-European (%)					2 (40)	7 (37)	5 (31)	5 (71)	4 (25)	9 (45)	1 (5)	33 (31)
Scores												
Mean score (%)	72	63	64	62	59	64	61	70	62	64	63	59
Mean score of participants passed (%)	74	74	72	72	70	75	73	76	72	73	74	73

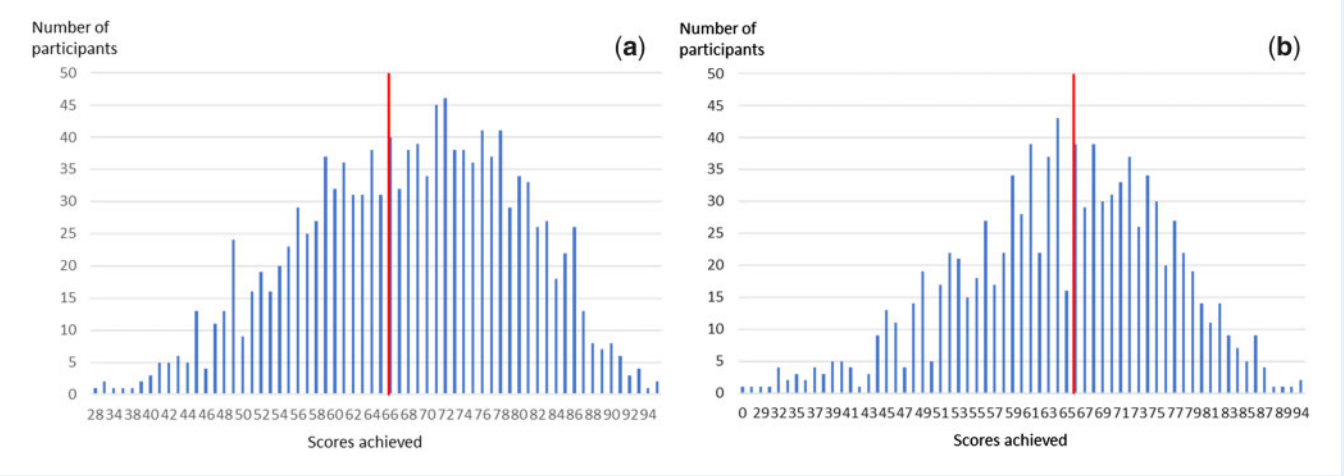

Figure 2 Frequency of scores achieved in exams: a pass threshold of 66% was defined since the first exam. (a) Clinical embryologist exams (2009–2018). **(b)** Senior clinical embryologist exams (2008–2018).

Table II The impact of candidate's academic degree on success of exam for Clinical and Senior Clinical Embryologists (data from 2013 to 2018).

	Clinical Embryologists[a]		Senior Clinical Embryologists[b]	
	Participants, *N*	Passed, *N* (%)	Participants, *N*	Passed, *N* (%)
BSc	302	146 (48.3)		
MSc[c]	437	239 (54.7)	379	184 (48.5)
PhD[d]	107	72 (67.3)	168	70 (41.7)
Total	846	457 (54)	547	254 (46.4)

The χ^2 test:
[a]Significant difference in pass rates between MSc and PhD ($P = 0.022$) and between BSc and PhD ($P < 0.001$) in clinical embryologist group.
[b]No difference in pass rates between MSc and PhD degrees in Senior Clinical Embryologist group.
[c]No difference in pass rates between MSc degrees of Clinical and Senior Clinical Embryologist exams.
[d]Significant difference in pass rates between PhD degrees in Clinical and Senior Clinical Embryologist exams ($P < 0.001$).

decreased. The questions in which distractors were chosen more often than the correct answer were considered as more difficult questions. The Clinical Embryologist exams included 6.9% and the Senior Clinical Embryologist exams 11.7% of such questions (data not shown).

The reliability index KR20 was always higher than 0.80 in all exams, showing a high level of internal consistency and that the exam produced similar results over subsequent years.

Discussion

Clinical Embryology has rapidly developed in the last three decades as a new discipline differentiated from medical embryology or the broader developmental biology. Specializations in laboratory techniques and skills required in human reproductive embryology now represent a new clinical laboratory discipline and, as such, have developed as a post-graduate educational competence. In 2013, most of Europe's 1700 ART laboratories were led by clinically oriented scientists, named 'Clinical Embryologists', who worked closely with gynaecologists to provide MAR (Kovačič *et al.*, 2015). In 2016, ESHRE presented the status of Clinical (Reproductive) Embryology and Clinical Embryologists in European countries to the UEMS, who recognized the ESHRE certification in Clinical Embryology and appraised the Clinical Embryologist and the Senior Clinical Embryologist exams.

Participants' profile

The ESHRE certification programme started 30 years after the birth of the world's first IVF baby and had to adapt to the existing highly heterogeneous staffing situation in ART laboratories, which has developed over decades. Staff composition, level of education, experience, etc. often depend on the size of MAR centres. Only in a few European countries are the educational standards and competences in ART laboratories defined in quality guidelines, national recommendations and/or legislations (Lamb, 2005; Practice Committees of ASRM *et al.*, 2014; Alpha Scientists in Reproductive Medicine, 2015; Go, 2015; Kovačič *et al.*, 2015; ESHRE Guideline Group on Good Practice in IVF Labs, 2016). Scientists from the natural/life sciences remain the prevalent staff in ART laboratories and up to 2008 no specific certification programme for Clinical Embryologists existed across Europe. One clear advantage of the ESHRE certification programme for Clinical Embryologists is that the exams are open to all candidates with a background in natural/life sciences and with relevant laboratory expertise and postgraduate practical experience in the field of MAR. From the collected data about exam participants' education, biologists strongly predominate, followed by biotechnologists and biomedical scientists.

The educational background of personnel in ART laboratories differs between European countries. Laboratory technicians/technologists predominate in some countries, while in others only biologists and

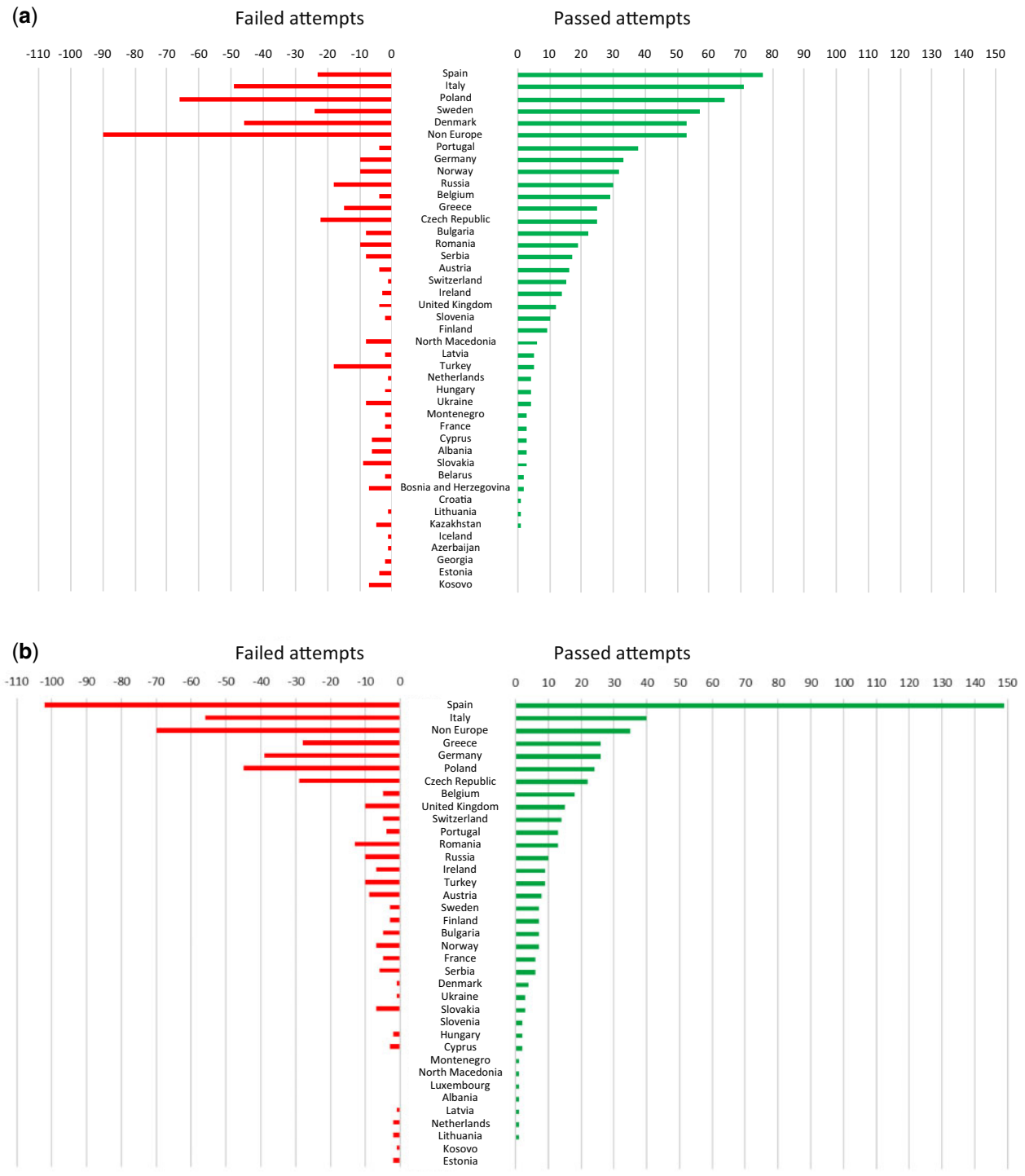

Figure 3 Number of failed and pass attempts of exams by country. (a) Clinical Embryologists (2009–2018). (**b**) Senior Clinical Embryologists (2008–2018).

medical doctors are allowed to work with human embryos (Kovačič *et al.*, 2015). Such differences could consequently influence the exam performance of candidates from specific countries, since the analysis of the performance per country and educational level shows that candidates with higher degrees (MSc and PhD) passed the exam at higher rates than those with a BSc only.

The ESHRE certification exams are typical criterion-referenced tests, where the main goal is for candidates to demonstrate a level of

Table III Post-exam analysis (2011–2018).

	Exams for Clinical Embryologists								
Year	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**Total**
Difficulty									
Mean difficulty index (p-statistic)	0.72	0.66	0.69	0.65	0.71	0.63	0.65	0.65	0.67
±SD	0.19	0.20	0.18	0.20	0.16	0.20	0.18	0.17	0.19
Median difficulty index	0.75	0.68	0.71	0.67	0.72	0.62	0.68	0.66	0.68
MCQs with optimal difficulty index (%)	60	71	67	73	72	81	79	82	73
Discrimination									
Mean discrimination index	0.27	0.23	0.27	0.25	0.29	0.28	0.28	0.28	0.27
±SD	0.12	0.11	0.11	0.12	0.15	0.13	0.12	0.12	0.12
Median discrimination index	0.28	0.23	0.27	0.25	0.28	0.31	0.29	0.28	0.27
MCQs with optimal discrimination index (%)	74	63	75	71	73	66	67	71	70
Distractor analysis									
Optimal distractors (%)	57	55	65	62	63	81	73	70	65
MCQs containing non-distractors (%)	23	26	18	16	10	4	7	8	14
Internal consistency									
Reliability index (KR-20)	0.87	0.82	0.86	0.85	0.89	0.87	0.87	0.88	0.86
	Exams for Senior Clinical Embryologists								
Year	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**Total**
Difficulty									
Mean difficulty index (p-statistic)	0.62	0.59	0.64	0.61	0.70	0.62	0.64	0.63	0.63
±SD	0.23	0.21	0.18	0.21	0.17	0.22	0.18	0.17	0.20
Median difficulty index	0.65	0.61	0.65	0.64	0.74	0.61	0.65	0.63	0.65
MCQs with optimal difficulty index (%)	70	81	77	79	67	72	84	84	77
Discrimination									
Mean discrimination index	0.25	0.27	0.27	0.26	0.27	0.25	0.29	0.24	0.26
±SD	0.15	0.14	0.13	0.12	0.14	0.15	0.15	0.12	0.14
Median discrimination index	0.27	0.29	0.27	0.26	0.28	0.26	0.28	0.26	0.27
MCQs with optimal discrimination index (%)	65	72	70	69	73	63	69	62	68
Distractor analysis									
Optimal distractors (%)	66	53	67	66	62	75	72	70	66
MCQs containing non-distractors (%)	27	16	12	20	22	8	13	11	16
Internal consistency									
Reliability index (KR-20)	0.84	0.87	0.87	0.85	0.87	0.82	0.88	0.84	0.85

Optimal difficulty index ranges from 0.2 to 0.8; optimal discrimination index ≥0.2; optimal reliability index ≥0.6; optimal distractors: answers chosen by ≥5% of examinees; non-distractors: not chosen answers. KR-20—Kuder–Richardson 20 is a reliability index, a special case of Cronbach's α.

practical experience and theoretical knowledge. The purpose of introducing two-level exams was to ensure that also holders of a BSc, who acquired the skills to perform ART laboratory techniques through work experience, could access the Clinical Embryologist certificate. The field of Clinical Embryology remains poorly represented in the university curricula (Hamilton and Carachi, 2014). To gain sufficient competence in practical tasks in an ART laboratory takes 3–4 years (Hughes and Association of clinical embryologists, 2012). According to European standards, to gain advanced knowledge of ART laboratory techniques, management experience and the ability to actively participate in clinical consultations, training usually takes a minimum of 6 years, which is similar to many medical specialities (ESHRE Guideline Group on Good Practice in IVF Labs 2016).

During the initial period of the ESHRE certification, concern was raised that the certificate only assesses theoretical knowledge and does not demonstrate the candidate's practical skills. Although this last point is highly relevant, formal practical training programmes for Clinical Embryology are largely non-existent at national level, with some valuable exceptions like the UK. In general, practical training programmes are left to the institutions themselves. They cover the prescribed programme that should be carried out in certified learning centres under the supervision of formal instructors using ethically

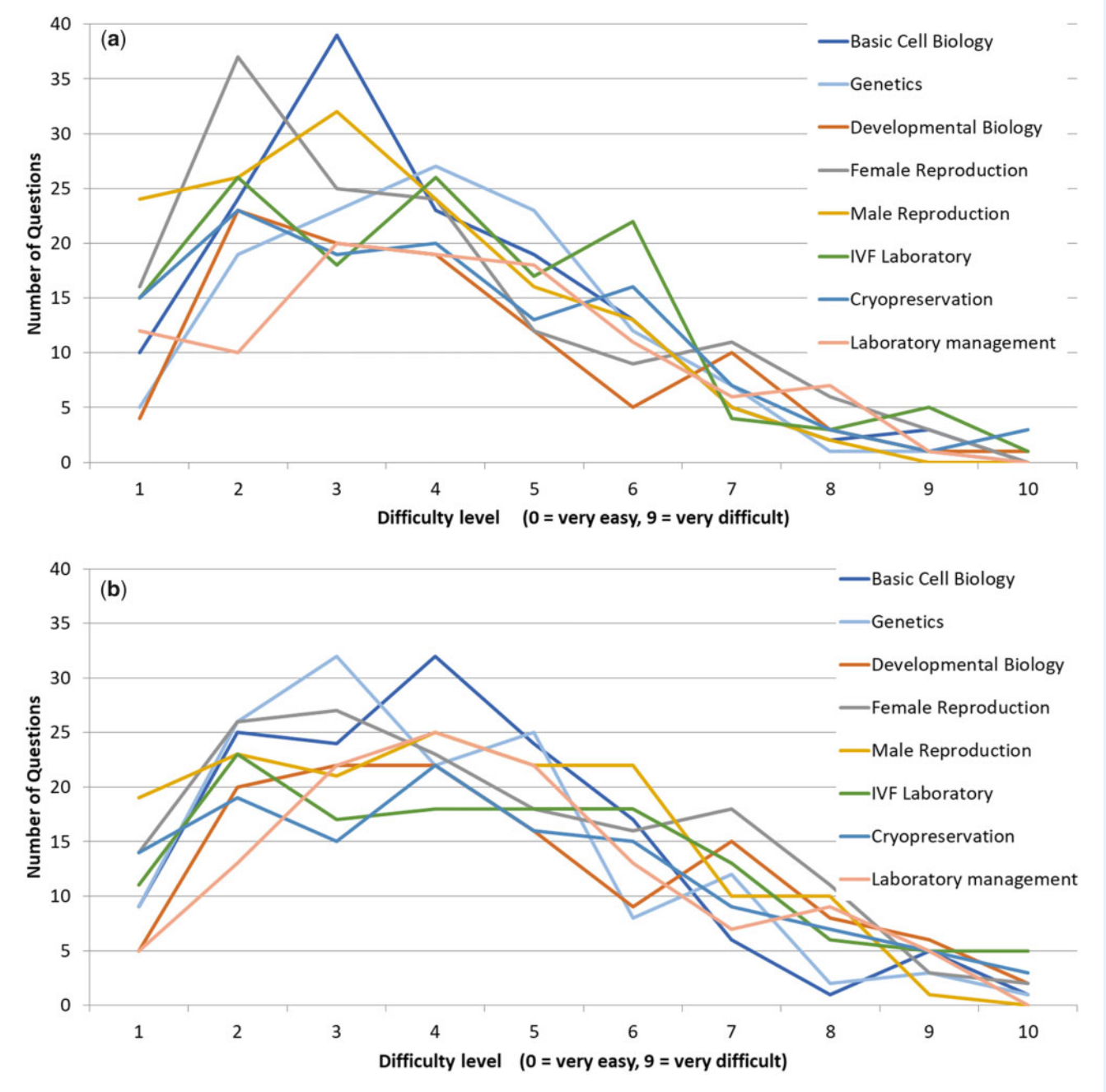

Figure 4 Difficulty level of exam questions by curriculum topic. (a) Clinical Embryologist exams (2009–2019). **(b)** Senior Clinical Embryologist exams (2008–2019).

acceptable biological material. The practical training must be recorded in predefined logbooks and qualifications through the verification of knowledge and issue of a certificate. Most European medical specialist exams operate in line with this principle and consist of theoretical and practical work.

A completed ESHRE logbook, signed by the supervisor, is compulsory and serves as confirmation that the candidate has been active in the laboratory for at least 3 or 6 years according to the proposed level, and has performed at least 50 procedures of each of the nine basic laboratory MAR working fields. While it is acknowledged that this is not a guarantee that the candidate has mastered the practical ART laboratory work, the completed logbook represents only a minimal proficiency entry criterion for acceptance to take the exam.

On average, 8% of applications were rejected. The most common reasons were an inadequate level of education, an incomplete logbook and/or inadequate working experience in an ART laboratory.

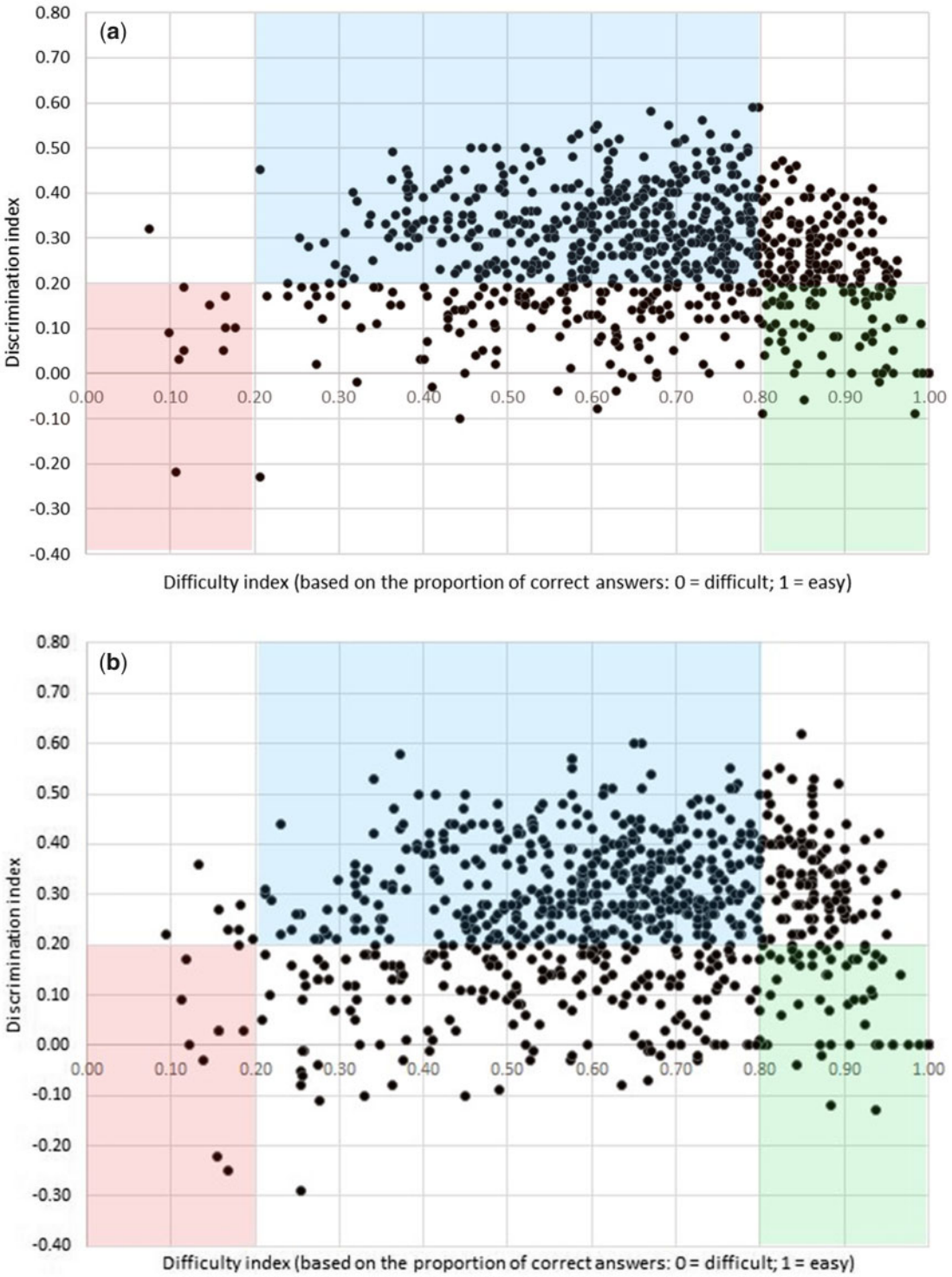

Figure 5 Post-exam analysis of multiple-choice questions. Scatterplot of difficulty and discrimination indices. Blue area: optimal discrimination and difficulty ranges. Pink area: suboptimal ranges (very difficult, low discrimination). Green area: suboptimal ranges (very easy, low discrimination). **(a)** Clinical Embryologist exams (2011–2018). **(b)** Senior Clinical Embryologist exams (2011–2018).

Exam questions background

MCQs with only one SBA from four answer options have been shown in previous analyses of medical specialist exams to be appropriate and objective. They allow for quick and simple manual or computer-based verification of the answers (Tenore *et al.*, 2015). In comparison with the false-true question type, where there is a 50% chance of obtaining the correct answer by guessing, MCQs give more reliable results (Mathysen *et al.*, 2013). Although MCQs also allow for guessing (Skakun *et al.*, 1979), this leads to lower results since one correct option (key) and three wrong options (distractors) decrease the chance for correct answer by guessing to only 25%. In order to prepare high quality questions with the optimal type of distractors, this type of question not only requires a good level of knowledge from the experts preparing the questions but also extensive discussions in order to attain the desired standards. The plausibility of distractors is a prerequisite for good MCQ building. Constructing plausible distractors represents one of the most difficult tasks in preparation of MCQs.

There are diverse views on what makes a high quality exam question. Assembling good questions is a task that needs to be developed over the years by the exam board members. In recent years, the EmCC has followed the guidelines for the composition of exam questions (The Australian Medical Assessment Collaboration, 2014; Tenore *et al.*, 2015), and members have attended workshops dedicated to this issue, organized by CESMA-UEMS. This has provided reassurance that the question structure meets the highest standards and has allowed for continual improvement in exam quality.

Inclusion of low quality questions can reduce overall exam reliability and validity. Problems may arise for the following reasons: the leading statement/question may be poorly written causing candidates to be confused; there may not be a clear correct answer, and one or more distractors could potentially qualify as the correct answer; there may be distractors that most candidates can understand as obviously wrong, increasing the odds of selecting the correct answer; inserted figures may not be clearly depicted (The Australian Medical Assessment Collaboration, 2014).

In order to minimize most of these issues in ESHRE Clinical Embryology exams, proactive measures included shortening the question length and ensuring that all answers were of similar length. In addition, all abbreviations or uncommon English words were removed.

However, there may be some additional factors that can add to the MCQ difficulty, which are unavoidable. For example, a single MCQ may test knowledge from more than one area of the curriculum, such as a MCQ testing a combined knowledge of Male Reproduction, Genetics and Developmental Biology. This could convey difficulty to those candidates who only work in specific areas of Clinical Embryology, and/or did not study the whole curriculum content. Furthermore, candidates from non-European countries could experience difficulty related to specific questions regarding EU legislation.

Given that the exam in the past 7 years was only in English, non-English-speaking participants may have been at a disadvantage. One might postulate that some Dutch, Spanish and Italian-speaking participants in the first three exams (provided in their native languages) may have had a better chance and that this may have resulted in a higher pass rate (Table I).

However, all of these factors were deliberate, aimed at approaching interdisciplinarity of curricular areas, addressing the EU legislation and using English as ESHRE's official language.

Number of applicants and pass rates

While in most undergraduate and university exams, the pass threshold is set *a priori* at 50% (10/20), in the postgraduate and professional exams, the pass threshold may also be set *a posteriori* at variable levels after each exam, based on post-exam analysis, such as by using the Angoff method (Tenore *et al.*, 2015).

In the ESHRE Clinical Embryology exams a pass threshold of 66% was set prior to the first exam, meaning that candidates had to select the correct answer for two-thirds of the questions in order to pass. Candidates were notified of this threshold prior to the exam, and this has been maintained for all exams.

Such a pass threshold, although arbitrarily set *a priori*, was chosen to ensure that only well-prepared candidates succeed in our Clinical Embryology exams. Furthermore, a detailed post-exam analysis has always been performed each year, allowing correction for eventual variabilities linked to a specific population of candidates and allowing results to be comparable to those obtained through *a posteriori* procedures, such as the Angoff method.

The analysis of applications revealed a great variability in the motivation for certification, number of attempts and the pass rates per country. It could be hypothesized that the number of applications per country depends both on the number of ART laboratory staff involved in national MAR programmes and recognition of the ESHRE Clinical Embryologist certificate by national professional bodies.

Several factors, such as the educational, legal or economical situation for each specific country, can play important roles in pass rates. Of relevance could be the access to scientific literature, meetings or research, and lack of experience with some techniques due to national legislation or clinic policy. For example, some clinics do not offer conventional insemination for IVF (preferring ICSI), preimplantation genetic testing (PGT) or cryopreservation. The lack of experience with the use of European standards and legislation for MAR centres (European Tissue and Cells Directives) for embryologists from non-EU countries might be one of several explanations for significant differences in pass rates between European and non-European candidates, respectively, in Clinical Embryologist (63% vs 33%) and Senior Clinical Embryologist (52% vs 31%) exams. However, the exam passing rate per country might be best interpreted by the national or regional professional bodies.

Another significant difference in pass rates was observed between first time and resit candidates, respectively, who participated in the Clinical Embryologist exams (63% vs 49%) and the Senior Clinical Embryologist exams (50% vs 39%). Initially candidates who failed the exam were allowed to re-apply to sit the next exam in the following year and the number of resits were increasing with years. However, a rule was introduced in 2016 to allow a maximum of three consecutive attempts. If the candidate failed on the third attempt, this person would have to miss the following year's exam, but would be eligible to apply for the exam in 2 years' time. This rule was intended to allow the candidate further time to study for the exam, and excluded candidates who might gamble on success due to the MCQ format. By freeing up this candidate space, it also allowed other candidates the opportunity to take the exam.

MCQ quality assessment

Preparing exams to a consistent level of quality is challenging. If the exam is considered 'too easy', this could lead to degradation of the

hard-earned reputation of ESHRE certification. On the contrary, if the exam is considered 'too difficult', this could reduce candidate motivation.

When preparing new questions, the difficulty level of a question cannot be determined until it is used. This is well known by psychometric science and for this reason specific statistical methods for monitoring the questions after use have been developed (Kuder and Richardson, 1937; Cronbach and Warrington, 1951 ; Ebel, 1954). Such feedback is also important for the group setting the questions.

The number of candidates participating in the exam (>250 per year) has become sufficiently large to allow reliable statistical analysis of exam performance, including question-response analysis (Kehoe, 1995). Such systematically performed analysis after each exam has allowed for continuous progressive improvement throughout the years. This has allowed for appropriate support in identifying those curriculum topics which need better or more detailed description for future exams.

Difficulty index

The exam difficulty level should help to correctly differentiate candidates regarding curriculum knowledge unless the MCQs are out of the optimal range of difficulty. An MCQ demonstrates low discrimination if it is either too difficult or too easy. The acceptable range of MCQ difficulty indices is usually arbitrarily defined, but this also depends on the question type and on the number and specific content of possible distractors (Crocker and Algina, 1986).

The optimal mean difficulty index for a 4-response MCQ is around 0.7, which is defined as 70% candidates selecting the correct answer (Oermann and Gaberson, 2017), but the optimal difficulty index range is 0.3–0.7 (The Australian Medical Assessment Collaboration, 2014; Oermann and Gaberson, 2017) or even larger, at 0.2–0.8 (Dixon, 1994). MCQs with difficulty indices outside the acceptable ranges need attention and should either be removed from subsequent exams or revised.

The ESHRE MCQs had a mean difficulty index of 0.7 and 0.6 in Clinical Embryologist and Senior Clinical Embryologist exams, respectively, with the majority of them in ranges of 0.3 to 0.95 and 0.2 to 0.95, respectively. The distribution of the difficulty indices of MCQs from all curriculum topics were spread across the entire acceptable range and very close to the desired range, but showed a deviation from normal distribution (right skewed) towards an easier area. Overall, this is a further characteristic demonstrating the reliability of the exams.

However, the upper range limit of 0.95 shows that the exams have included several easy MCQs. These could subsequently be revised for future exams to bring them into the acceptable range (Dixon, 1994; The Australian Medical Assessment Collaboration, 2014; Oermann and Gaberson, 2017). The mean pass rates of the Clinical Embryologist and Senior Clinical Embryologist exams, 60% and 50%, respectively, demonstrated that these exams have achieved a status of moderately difficult to difficult exams. Difficulty indices of most of the MCQs were considered optimal, with 73% and 77% of the Clinical Embryologist and Senior Clinical Embryologist MCQs, respectively, in the range between 0.2 and 0.8 (Dixon, 1994).

Inclusion of other types of question types, such as MCQs of the SBA type with five answer options instead of four, MCQs with several possible answers instead of just one, and extended matching questions may provide an increased assessment of knowledge. All these new question types may reduce the chance of getting the right answer by guessing. However, there are also disadvantages with their use, and further discussion will be required.

Discrimination index

Reviewing the answers to each individual MCQ by the best and the worst exam performers has provided information on the discriminating potential of the question. If a MCQ had a low difficulty index but a high discrimination index, this meant that the MCQ was difficult but relevant. MCQs that demonstrated good discriminating potential were usually moderately difficult, while very easy or very difficult MCQs usually had poor or even negative discrimination potential (Sim and Rasiah, 2006). Where candidates answered a MCQ correctly but scored low in the exam overall, or they answered an MCQ incorrectly but scored high in the exam overall, both scenarios describe a low discrimination index. If a MCQ had a negative discrimination index, this usually indicated that the MCQ required removal or correction for future use. The exam analysis showed that only 1.8–3.8% of MCQs had a negative discrimination index. Altogether, 68–70% of MCQs passed the criteria for good discrimination.

Distractor analysis

An optimal distractor must look like a correct answer. It is considered a non-distractor if it is only selected by <5% candidates (Haladyna and Downing, 1993). A trend for improving the quality of MCQs was achieved for both exam levels over several years by increasing the proportion of optimal distractors. By increasing the number of answer options per MCQ from 4 to 5, the negative effect of one non-distractor on the quality of MCQ is minimized.

Reliability coefficient KR-20

The KR-20 is a coefficient for the assessment of exam quality and represents an estimation of its internal consistency. The KR-20 depends on the difficulty, the total number of MCQs and the number of MCQs that did not discriminate (Kuder and Richardson 1937). A low KR-20 indicates that many MCQs are too difficult, too easy, or poorly written (McGahee and Ball, 2009). All ESHRE Clinical Embryology exams had a KR-20 in the range of 0.8–0.9 and were therefore considered as very good to almost the best standardized tests, according to previously described criteria (Nunnally, 1967).

Curriculum analysis

The MCQs were divided in eight curriculum topics; however, there is inevitably some overlap. For example, the topic of Basic Cell Biology MCQs may contain questions on meiosis, which overlap the topic of Genetics.

In addition, certain curriculum topics have changed significantly over time. For example, the topic of Cryopreservation has undergone substantial change with the introduction of vitrification. Furthermore, laboratory techniques of embryo biopsy and PGT, as well as the views on the relevance of their use, have changed considerably over the last 10 years. Consequently, some MCQs were removed, modified or newly introduced in order to keep up with the new trends in our field.

The new ESHRE guidelines, the Atlas of Human Embryology, and consensus publications (e.g. Magli *et al.*, 2012; ESHRE Guideline Group on Good Practice in IVF Labs2016 ; ESHRE SIG-E and Alpha, 2017) has resulted in the introduction of new questions derived from these documents. Attempts were made to include schemes and photo material in the exam, but due to the altered quality of the photographs on the printed sheets, this practice did not expand as anticipated. The use of electronic exams in the coming years will facilitate and possibly expand this possibility. Depending on the specific modifications, the list of recommended literature was regularly and accordingly updated. It is reassuring that, despite being 10 years old, the curriculum still conveniently covers all major areas. However, it may benefit from updates in the near future.

Over this 10-year period, various questions have been asked by candidates about the application process and the exams. To assist with this, a 'Frequently Asked Questions' section was added the ESHRE website, which is updated as required. Complaints have been marginal, with no complaints received from candidates who passed the exams.

Conclusion

Both the Clinical Embryologist and the Senior Clinical Embryologist exams are demanding and have adequate discrimination power. The difficulty grade is evenly distributed across the eight topics of the curriculum. Furthermore, the exams demonstrated high reliability and allowed for appropriate scrutiny of the knowledge of Clinical Embryologists.

The interest and performance in the ESHRE Clinical Embryology exams varies among candidates from different countries. Academic level and educational background were factors associated with differences in performance.

The ESHRE exams on Clinical Embryology are the most widely accepted tests of knowledge from laboratory science in MAR. These exams provide the recognition that minimum standards have been achieved. The acquired certificate allows ESHRE certified Clinical and Senior Clinical Embryologists to be included in the ESHRE CPD programme and registered on the list of ESHRE certified and CPD active Clinical Embryologists.

Clinical Embryology is increasingly recognized as a specific discipline for scientific staff collaborating closely with clinicians in managing human infertility by MAR. ESHRE Certification for Clinical Embryologists, and its current recognition by CESMA-UEMS, adds clarification to this health-staff postgraduate examination, and conveniently adds Clinical Embryology into the wider group of health disciplines harmonized through professional societies such as ESHRE and EBCOG.

The EmCC will continue its efforts to reinforce ESHRE's position at UEMS-CESMA and to follow its advice. The authors hope that the certification results shown will enhance recommendations towards delivering learning programmes of knowledge in Clinical Embryology and will increase the interest for ESHRE certification programmes worldwide.

Supplementary data

Supplementary data are available at *Human Reproduction Open* online.

Acknowledgements

The authors thank Kay Elder for providing initial pool of MCQs and to all past EmCC members: Joyce Harper, Etienne van den Abbeel, Julius Hreinsson, Maria José de los Santos, Giovanni Coticchio and Heidi Rojiemans for their valuable contribution to ESHRE Clinical Embryologist certification programme. For the pilot exam in India, we thank Kuldeep Jain and Jayant Mehta from the Indian Fertility Society. Thanks also to Ioannis Messinis for critical reading of the manuscript, to Wil van Groesen for development of on-line application platform and to Žiga Kovačič for his contribution to the statistical analysis.

Authors' roles

B.K. and C.E.P. co-ordinated the analysis and writing the manuscript. F.J.P., as MCQ bank manager, provided raw data and methodological support for further data analysis. C.P. provided administrative and technical support in preparation and organization of exams and their evaluations. All other authors contributed to the interpretation of the analysis and critically revised the manuscript.

Funding

The study has no external funding.

Conflict of interest

The authors reported no conflicts of interest.

References

Alpha Scientists in Reproductive Medicine. The Alpha Consensus Meeting on the professional status of the clinical embryologist: proceedings of an expert meeting. *Reprod Biomed Online* 2015;**30**:451–461.

Calhaz-Jorge C, Feki A, Farquharson R. European view of subspecialty training on behalf of the European Society of Human Reproduction and Embryology (ESHRE). *Fertil Steril* 2015;**104**:8–11.

Crocker L, Algina J. *Introduction to Classical and Modern Test Theory*. New York: Holt, Rinehart and Winsto, 1986.

Cronbach LJ. Warrington WG. Time-limit tests: estimating their reliability and degree of speeding. *Psychometrika* 1951;**16**:167. 10.1007/BF02289113 14844557

Dixon RA. Evaluating and improving multiple choice papers: true–false questions in public health medicine. *Med Educ* 1994;**28**:400–408.

Downing SM, Baranowski RA, Grosso LJ, Norcini JJ. Item type and cognitive ability measured: the validity evidence for multiple true-false items in medical specialty certification. *Appl Measur Educ* 1995;**8**:187–197.

Ebel RL. Procedures for the analysis of classroom tests. *Educ Psychol Meas* 1954;**14**:352–364.

ESHRE Special Interest Group of Embryology and Alpha Scientists in Reproductive Medicine. The Vienna consensus: report of an expert meeting on the development of ART laboratory performance indicators. *Reprod Biomed Online* 2017;**35**:494–510.

ESHRE Guideline Group on Good Practice in IVF Labs, De los Santos MJ, Apter S, Coticchio G, Debrock S, Lundin K, Plancha CE, Prados F, Rienzi L, Verheyen G, Woodward B, Vermeulen N. Revised guidelines for good practice in IVF laboratories (2015). *Hum Reprod* 2016;**31**:685–686.

Go KJ. 'By the work, one knows the workman': the practice and profession of the embryologist and its translation to quality in the embryology laboratory. *Reprod Biomed Online* 2015;**31**:449–458.

Haladyna TM, Downing SM. How many options is enough for a multiple-choice test item? *Educ Psychol Meas* 1993;**53**:999–1010.

Hamilton J, Carachi R. Clinical embryology: is there still a place in medical schools today? *Scot Med J* 2014;**59**:188–192.

Hughes C, Association of clinical embryologists. Association of clinical embryologists—guidelines on good practice in clinical embryology laboratories 2012. *Hum Fertil (Camb)* 2012;**15**:174–189.

Kehoe J. Basic item analysis for multiple choice tests. *Pract Assess J Res Eval* 1995;**4**:10.

Kelley TL. The selection of upper and lower groups for the validation of test Items. *J Educ Psychol* 1939;**30**:17–24.

Kovačič B, Plas C, Woodward BJ, Verheyen G, Prados FJ, Hreinsson J, De Los Santos MJ, Magli MC, Lundin K, Plancha CE. The educational and professional status of clinical embryology and clinical embryologists in Europe. *Hum Reprod* 2015;**30**:1755–1762.

Kuder GF, Richardson MW. The theory of the estimation of test reliability. *Psychometrika* 1937;**2**:151–160.

Lamb VK. National occupational standards in clinical embryology: implementation in 2005? *Hum Fertil* 2005;**8**:1–4.

Magli MC, Jones GM, Lundin K, van den Abbeel E. Atlas of human embryology: from oocytes to preimplantation embryos. Preface. *Hum Reprod* 2012;**27**(Suppl 1):i1.

Mathysen DG, Aclimandos W, Roelant E, Wouters K, Creuzot-Garcher C, Ringens PJ, Hawlina M, Tassignon MJ. Evaluation of adding item-response theory analysis for evaluation of the European Board of Ophthalmology Diploma examination. *Acta Ophthalmol* 2013;**91**:e573–e577.

McGahee Ball TW, J. How to read and really use an item analysis. *Nurse Educ* 2009;**34**:166–171.

Nunnally JC. *Psychometric Theory*. New York: McGraw-Hill, 1967, pp. 172–235.

Oermann M, Gaberson K (eds). *Evaluation and Testing in Nursing Education*. New York, Springer Publishing Company, 2017, p. 403.

Practice Committee of the American Society for Reproductive Medicine; Practice Committee of the Society for Assisted Reproductive Technology; Practice Committee of the Society of Reproductive Biologists and Technologists. Recommended practices for the management of embryology, andrology, and endocrinology laboratories: a committee opinion. *Fertil Steril* 2014;**102**:960–963.

Richardson MW, Stalnaker JM. A note on the use of bi-serial r in test research. *J Gen Psychol* 1933;**8**:463–465.

Sim SM, Rasiah RI. Relationship between item difficulty and discrimination indices in true/false multiple choice questions of a paraclinical multidisciplinary paper. *Ann Acad Med Singapore* 2006;**35**:67–71.

Skakun EN, Nanson EM, Kling S, Taylor WC. A preliminary investigation of three types of multiple choice questions. *Med Educ* 1979;**13**:91–96.

Tenore A, Mathysen DGP, Mills P, Westwood M, Rouffet J, Papalois V, Sparrow O, Goldik Z. A guide to successfully writing MCQs: Executive summary. *UEMS-CESMA Publication* 2015.

The Australian Medical Assessment Collaboration (AMAC). Determining the Quality of Assessment Items in Collaborations: Aspects to Discuss to Reach Agreement. Australian Government Department of Education, Canberra, Australia 2014, 26 p.

Trávník P, Hampl A, Hüttelová R, Malenovská A, Priesnitz J, Rejthar D, Žáková J. Theoretic and practical content of the clinical Embryology. *Ceska Gynekol* 2013;**78**:400–401.