

# An Alternative Certification Examination (“ACE”) in Surgery<sup>☆</sup>

Marie Catherine Morris, MSc, BSc (Hons), RGN,<sup>\*</sup> Amy E. Gillis, MCh, BSc, SSO Accredited, RCPSC Accredited,<sup>†</sup> Craig O. Smoothery, MSc, MBA, BSc (Hons), BA,<sup>‡</sup> Martina Hennessy, PhD, MB, BA, RGN, Dip. Toxicology (Cardiff),<sup>\*</sup> Kevin C. Conlon, MA, MCh, MBA, FRCSI, FACS, FRCS, FTCD,<sup>\*†</sup> and Paul F. Ridgway, MD, MMedSc, FRCS (Gen Surg)<sup>\*†</sup>

<sup>\*</sup>Education Division, School of Medicine, Trinity College, University of Dublin, Dublin, Ireland; <sup>†</sup>Department of Surgery, Trinity College, University of Dublin, Tallaght Hospital Campus, Dublin, Ireland; and <sup>‡</sup>School of Electrical, Electronic and Mechanical Engineering, University College Dublin, Belfield, Dublin, Ireland

**BACKGROUND:** Medical graduates are required to be competent in many domains of professional practice when joining the health care workforce. Current undergraduate examination methods robustly assess up to 5 of these 8 required skills. This study sought to evaluate an alternative certification examination (“ACE”) in assessing all of the 8 required domains in surgical cases.

**METHODS:** A total of 143 final-year medical students were invited to participate in an “ACE” in February 2013. In total, 137 students, 95.8% of the class, agreed to participate. The “ACE” format consisted of 4 sequential patient encounters observed by 2 independent examiners. It assessed all the 8 required domains of professional practice. The examiners and the students evaluated this examination format using a Likert scale and free-text comments.

**RESULTS:** The “ACE” assessed all the 8 domains. The inclusion of a patient safety measure to avoid an egregious error in the pass criteria resulted in 27 (18.9%) students failing to meet them. The correlation of grades between the independent examiners in the “ACE” was strong at a Cronbach  $\alpha$  of 0.907 (CI: 0.766–1). The “ACE” format was reported as an acceptable examination methodology by the examiners for formative or summative assessment of surgical cases at the end of a primary medical degree.

**CONCLUSION:** The “ACE” format is standardized, is integrative, and has excellent interrater reliability. Inclusion of a patient safety measure as pass criteria appears to increase specificity. The “ACE” shows potential as an alternative examination to the traditional long case examination and

objective structured clinical examination in assessing all the 8 domains of professional practice. (J Surg ■■■■-■■■. ©2014 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

**KEY WORDS:** assessment, surgery, patient safety, competence

**COMPETENCIES:** Patient Care, Medical Knowledge, Practice Based Learning and Improvement, Systems Based Practice, Professionalism and Interpersonal Skills and Communication

## INTRODUCTION

Undergraduate assessment is a prerequisite to licensure, culminating in the awarding of a basic qualification to practice medicine. Medical schools strive to ensure that students have acquired the expected knowledge, skills, and attitudes over a broad range of general and specialty medicine to ensure patient safety. Regulatory bodies have a statutory responsibility to set and maintain the standards of basic medical education during undergraduate training and internship. The European guidelines suggest that all medical graduates be competent in core clinical skills<sup>1</sup> to enable greater “freedom of movement” for graduates to work in Europe.<sup>2</sup>

One of the many requirements of medical schools is to ensure graduates are professionally competent and adequately prepared for joining the hospital workforce.<sup>3,4</sup> An assessment format that reliably measures the breadth of knowledge and clinical skills acquired over the entirety of the medical curriculum is paramount to ensure that only students that achieve the required grade pass. Graduates of medical training must be able to apply the knowledge that they have accumulated during their training. Ultimately,

<sup>\*</sup>The cost of running the examination was met by an Educational Grant from the School of Medicine, Trinity College Dublin, University of Dublin, Dublin 2, Ireland.  
Correspondence: Inquiries to Marie Catherine Morris, MSc, BSc (Hons), Education Division, Trinity College/Tallaght Hospital, University of Dublin, Dublin, Ireland; fax: (189) 6-2910; e-mail: morrism4@tcd.ie

they must demonstrate competence in the domains of professional practice namely clinical skills, communication skills, scholarship, professionalism, relating to patients, collaboration and teamwork, and management (including self), and at the core of all requirements is ensuring patient safety and quality of patients' care<sup>4</sup> (Fig. 1).

There are many methods of summatively assessing competence in these domains in medical students, with the most common methods including a long case examination or an objective structured clinical examination (OSCE) or both. Short case examinations may also be used in conjunction with the former methods. The long case examination is the traditional method of assessing clinical skills at the end of a primary degree and includes the assessment of history taking, physical examination, and data interpretation skills in an oral examination. Students are assessed in medical and surgical cases in this format and are expected to achieve a pass mark in both, with no compensation across specialties. Assessment is usually summative and a barrier to qualification. Worldwide, the long case examination is routinely used as a high-stakes examination.<sup>5</sup> A recent literature review on the long case examination yielded 18 articles<sup>6</sup> and concluded that its format may vary between institutions but generally involves students spending 30 to 45 minutes alone with a patient. During this time, they take a medical history and perform a physical examination to identify clinical signs of disease. Then, they present a summary of their findings to 2 or 3 subject experts who mark them on a global rating scale. A diagnosis and management plan is often sought. Students are generally examined on one or more medical and surgical cases, and the final grade is typically an average of all the examiner's findings. The reported primary

advantage of this assessment method is that it is holistic and authentic.<sup>6</sup> Many researchers report and caution that the long case examination attempts to generalize the abilities of a students across a broad spectrum of clinical scenarios with a single case<sup>7-10</sup>; thus, ideally a few cases should be examined.

To address the potential shortcoming of the traditional long case examination, modifications have been made in recent years, resulting in alternative examination formats. Wass and Jolly<sup>10</sup> found that observation of history taking in the long case examination constituted a distinct component of clinical competence, which is missed with the traditional approach of nonobservation and presentation of findings. Further modifications include the mini-clinical evaluation exercise,<sup>11</sup> objective structured long examination record,<sup>12</sup> and direct observation of clinical encounter examination.<sup>13</sup> In these modified long case examinations, students interviewed and examined 4 patients with 2 sets of examiners, observing with a reported generalizability coefficient of 0.84. Luiz et al.<sup>14</sup> subsequently found that when a student took 2 structured, standardized, observed long case examinations, each marked by a different examiner, that examiner's agreement in clinical skills assessment was very high at 89%.

With regard to the domains of professional practice the long case examination examines 5 of the 8 criteria but does not overtly assess professionalism, collaboration and teamwork, and patient safety.

The OSCE is another common method of assessing clinical skills in the early years of the medical degree program. Despite standardization of assessment of procedural skills with checklists and global raters, the demonstration of competence over a series of independent skills,



**FIGURE 1.** The 8 domains of good professional practice.

however, does not guarantee the ability of a student to manage an entire patient episode. The expected level of competence in the workplace today includes areas such as patient safety, communication, collaboration, teamwork, management, professionalism, and clinical skills.<sup>4</sup> The alternative certification examination (“ACE”) format, by examining an entire episode with a patient, has been shown to address the fragmentation of skills assessment using an integrative approach, albeit with trained actors.<sup>15</sup>

The authors have previously investigated the feasibility of using an integrative postgraduate examination format (patient assessment and management examination [PAME])<sup>16</sup> in an undergraduate setting. This new application was previously dubbed the “ACE.” Promising results were found in the assessment of surgical cases.<sup>15</sup> The “ACE” format proved feasible and discriminatory. It was deemed more integrative than an OSCE by examiners, as it included direct observation of communication, time management, and resource management, as well as the core skills required to manage an entire patient encounter. The “ACE” format demonstrated a potential to complement current elements of the prepractice (intern/foundation year 1) clinical assessment barrier. This study further investigated the “ACE” format in assessing all of the 8 required domains of professional practice in surgical cases.

## METHODOLOGY

### Research Setting

This study occurred in a teaching hospital affiliated to a Dublin University Medical School.

### Research Aims

The research aims are given in [Table 1](#). See [Table 1](#) for a list of research aims.

### Ethics

Local ethics committee approval for this research was sought and received with the stipulation that the researchers were to ensure that the examiners used in the “ACE” would not be involved in the subsequent long case clinical

examination. This stipulation was adhered to. A written informed consent was gained from all the participants to use the data collected from “ACE” and long case examination for research and publication purposes.

### Research Sample

Following approval from the Health Sciences Division Ethics Committee, a final-year medical class ( $n = 143$ ) was invited to participate in a research project, which was an extra formative examination using the “ACE” format in February 2013. Participation was voluntary and students were assured that they could decline to participate and withdraw from the study at any time with no repercussions. The participating students were furnished with the “ACE” format 2 weeks in advance of the examination and were advised that their opinions of the examination format would be sought at a later date. Marking sheets were coded to ensure that the data remained anonymous. The students were advised that written feedback on their performance from all the examiners would be provided within a week of the examination. The option of meeting with the researchers to discuss any aspect of the grading was offered to all. A written informed consent was obtained, and specific agreement sought from students to use the data from the findings for research publication purposes. Anonymity and confidentiality were assured. Registrars and senior house officers were used as examiners in “ACE.” No consultants were used as this group of experts would examine these students 4 months later in their final clinical long case examinations, this was to eliminate the potential for a power relationship and remove potential bias. A total of 137 students agreed to participate in the study. This represents a response rate of 95.8%.

### Validity of Tools

Previously validated global assessment tools from the “PAME” methodology<sup>16</sup> were used as the basis for this research ([Appendix 1](#)). A pilot study using the “ACE” format was conducted the previous year to assess the feasibility of the “ACE” as an alternative or adjunctive examination format. The encounters, marking sheets, timings, case content, and complexity and evaluation forms were all tested and evaluated by the examiners before the study.

**TABLE 1.** Research Aims

Item	Research Aim
1	To assess all the 8 domains of professional practice in 1 examination of surgical cases.
2	To investigate the interrater reliability between the independent examiners in the “ACE” format.
3	To examine what effect, if any, the inclusion of a patient safety measure requiring the avoidance of an egregious error in the “ACE” format had on pass rates.
4	To investigate final-year medical students’ views on the “ACE” format in comparison with an OSCE assessment method in clinical skills.
5	To investigate the examiners’ views on the “ACE” format in comparison with a traditional long case assessment method in clinical skills.

## Reliability of Data

All examiners were practicing surgeons who would not be involved in the subsequent final long case examination in April. Everyone was furnished with the standardized marking sheets and examination format 2 weeks before the examination. A standard setting meeting was held on the morning of the examinations with standard instructions to all the examiners. An agreement was reached on pass marks and pass criteria. Two examiners assessed each candidate at each case across all 4 encounters and marked students independent of each other. All the examiners were familiar with the clinical case scenarios used. Five pairs of examiners were present each day. The examination ran over 2 days. The researcher (M.M.) remained on the examination floor to answer any queries with regard to the examination.

## Standardised Simulated Patients

Professional actors were used as standardized simulated patients, and everyone had been involved as role-playing patients in undergraduate and postgraduate medical clinical examinations previously. Written instructions on the clinical condition and standardized questions to ask the students were provided 2 weeks before the examination. A contact number was provided for any queries. A meeting was held on the morning of the examination to ensure standardization in the presentation of the clinical problem and in answering questions. Actors were also asked to note the “professionalism” of the students, e.g., their self-presentation, bedside manner, listening skills, perceived interest in them and their illness, and their confidence in the student’s ability to manage their condition effectively (Table 2). The researcher (M.M.) was available on the examination floor at all times to address any queries arising.

## Examination Format

All students assessed 1 of 2 standardized simulated patients both presenting with a common surgical complaint: right iliac fossa pain or right upper quadrant pain, across 4 sequential encounters, consisting of history taking,

examination, interpretation of test results, diagnosis and management, and finally a follow-up review encounter in consultation with their clinical team (Fig. 2). The total examination time was 29 minutes. The student entered an examination center and remained in that center with the patient and the 2 examiners for the entire 29 minutes. A bell was rung when the time for each individual encounter had passed. The bell signaled to the examiners that the students should move onto the next encounter within the same room.

## Pass Criteria Including Professionalism and a Patient Safety Measure

Competence was assessed in the following 3 ways within the “ACE” format (Appendix 1):

- (i) By the ability of the student to score at least 3/5 (barely adequate for intern/foundation year 1–level practice) in all the 4 encounters. The domains assessed were scholar, clinical skills, relating to patients, communication skills, and self-management including time and collaboration.
- (ii) By demonstrating the domain of professionalism—for the purpose of this study professionalism was defined as avoidance of all the behaviors listed in Table 2. This criterion was devised by expert consensus in the research group and was sent to the examiners and the actors 2 weeks before the examination, reiterated 1 week before, and was again discussed on the morning of the examination. A written copy of this table was supplied to the examiners and the actors for consultation during the examination if required. A yellow note was attached to the marking sheets to demonstrate concerns regarding professionalism, and the reasons for same were listed on the marking sheets under “comments.” This was a decision made by the examiner in collaboration with the simulated patient.
- (iii) By demonstrating the domain of patient safety and quality of patient care, i.e., by avoiding an egregious error when deciding differential diagnoses that

**TABLE 2.** Deficits in Professionalism

Item	Deficit in Professionalism	Examples
1	Inappropriate attire	Not complying with the college dress policy.
2	Poor rapport with the patient	Poor verbal and nonverbal communication skills.
3	Displays of disinterest	Poor eye contact, looking at notes, and failure to listen.
4	Lack of regard for patient’s anxiety, discomfort, or pain	Failure to look at facial expressions of pain and no display of sympathy or reassurance.
5	Inappropriate familiarity with the patient	Casual approach and inappropriate humor.
6	No attempt to involve the patient in the management plan	Paternalistic approach, no discussion with patient regarding plan, and no opportunity for asking questions.
7	Disregard of patient’s concerns	Dismissive approach, no evidence of reassurance, and no mention of seeking a senior or second opinion.

### Schema of "ACE"

Exam Centres – common surgical cases (0.5 hour per station); Observed by examiners/raters, standardised patient and materials (referral letters, imaging, investigation results); Timings approximate.

Standardised Patient Initial assessment

#### Encounter 1

History (6 mins)

Examination (6 mins)

#### Encounter 2

Order and Interpretation of investigations results (5mins)

#### Encounter 3

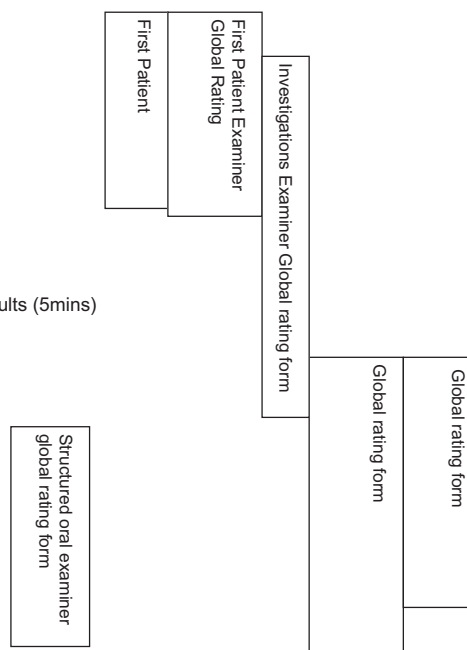
Diagnosis and management (6mins)

Structured Oral Examination

#### Encounter 4

Review of case (6 mins)

In consultation with Team



**FIGURE 2.** Schemata of "ACE" format.

could result in deterioration in the patient's condition during the patient's assessment and management.

In this examination, the following potential egregious errors were included:

- (a) Failing to consider a potential ectopic pregnancy in a female patient with right iliac fossa pain.
- (b) Failing to consider a potential perforated duodenal ulcer in a male patient presenting with right upper quadrant pain.

The rationale for the inclusion of a patient safety measure as pass criteria was to ensure that those achieving a pass mark were demonstrating competence in all the 8 domains as required by the certifying bodies.<sup>4</sup> Students who did not consider these life-threatening conditions, thus potentially compromising patient safety, had a red note applied to their marked sheets and were awarded a final score of 2/5, which was a fail. Both the examiners made this assessment independent of each other.

### Feedback to Participants on "ACE" Performance

All the students received a written copy of their coded marked assessment sheets with the examiners' comments

and were offered the opportunity to meet and discuss their performance with the researchers (M.M. and A.E.G.) within a week of the examination. The researchers attended the lecture hall to run through the pass/fail criteria and the significance of a yellow or red note on the marked sheets. All those who had a red or yellow note on their marked sheets were strongly advised to meet with the researchers so this could be discussed and support and remediation offered.

### Evaluation of "ACE" Format

All the students and the examiners completed an evaluation form on the "ACE" format (Appendix 2). The form used was devised by the research team and included a 5-point Likert scale and the opportunity for free-text comments. Face and content validity of this form was tested and retested previously by 10 senior medical staff not involved in this study. This process was repeated until there was an expert consensus that the form measured what the study was seeking to evaluate. The examiners were asked to compare the "ACE" with a long case format and an OSCE format. Students had not yet experienced a long case examination format at this stage of the course; hence, they were asked to compare the ACE with an OSCE format. Simulated patients were asked by the examiners for feedback on the

**TABLE 3.** Fail Grade Odds Ratio's

	Fail
Odds ratio	33.1
99% CI	(3.78–6374)

interpersonal, communication, and professional skills of the students at the end of each 29-minute examination.

## Data Analysis

Expert statistical advice was sought on data collection and analysis. Overall, 3 different interrater reliability (IRR) tests were performed to address any potential shortcomings of each individual test. These tests included the Cronbach alpha, Krippendorff alpha, and Cohen alpha.<sup>17</sup> An odds ratio was obtained to investigate the specificity of the “ACE” format. As data from the Likert scales in the evaluation forms were nonparametric, median score and interquartile ranges (IQR) are reported. Free-text comments from the examiners are included.

## Quantitative Results

### ACE: Analysis of Fail Grades

The proportion of students failing the “ACE” was compared with those failing a subsequent long case examination using odds ratio. In this instance, the effect size of the “fail” group in the “ACE” was 33.1 times that of the effect size on the subsequent long case examination (Table 3).

### Analysis of “ACE” Fail Group

A total of 27 students failed the “ACE.” The justification for failure (not avoiding an egregious error when considering differential diagnoses, low grade, or both) is broken down in Table 4. “Gross grade” means the grade without considering the patient safety aspect.

### IRR for the “ACE” Format

The examination data set was filtered as follows:

- Students having 2 examiners
- Examiner pairings that examined together for the entire examination

**TABLE 4.** “ACE” Fail Distribution

Considered Patient Safety	Gross “ACE” Grade		
	Fail	All Pass Grades	Total
No	13	0	13
Yes	8	6	14
Total	21	5	27

**TABLE 5.** Interrater Reliability of the “ACE” Format

Metric	Value	95% CI
Cronbach $\alpha$	0.907	0.766–1
Krippendorff $\alpha$	0.824	0.824–1
Cohen $\kappa$	0.823	0.571–1

The traditional first- and second-class honor grades awarded in the “ACE” were amalgamated as a “Pass.” Consequently, only 1 pass grade existed for the purpose of the analysis. There were 5 examiner pairings with each pairing examining between 12 and 14 students. Where examiners assessed alone for any period of the examination, these data were excluded from the analysis. Bootstrap techniques were used to estimate the 95th percentile CIs for each of the metrics<sup>17</sup> (Table 5).

## Qualitative Results

### The Examiners Evaluations

All the examiners completed an evaluation form using a 5-point Likert scale (Appendix 2). The responses to each statement were generally positive. The “ACE” format was perceived as a comprehensive assessment method, with a median score of 4 (IQR: 4–5). It was deemed strong at identifying students’ strengths and weaknesses, with a median score of 4 (IQR: 4–4). The “ACE” format was reported as superior when compared with the OSCE in assessing an entire patient episode, with a median of 4 (IQR: 4.5–5). Moreover, it was deemed more discriminatory than the long case format, with a median of 4 (IQR: 4–4).

On review of the free-text comments, examiners reported that the “ACE” format was useful in preparing students for managing a patient in the clinical setting. Examiners also reported that the time lines of 30 minutes per case proved adequate for the surgical cases examined. Examiners were positive regarding the examination format. They agreed that the integrative approach with case management clearly differentiated between good and weak students, reporting that “there was nowhere to hide.” Examiners reported that the “ACE” format clearly highlighted deficits in knowledge in history taking and examination and welcomed the addition of patient safety as passing criteria on safety to practice grounds. Moreover, the formal assessment of professionalism was seen as a very important factor, as often problems with regard to professional behavior are not tangible. Two (20%) examiners reported their preference for “real” patients with 8 (80%) reporting the benefit of standardization with actors. The examiners acknowledged the excellence and credibility of the professional actors and reported that logistically the “ACE” format was considerably easier to plan and deliver than the OSCE and long case formats.

## Participant Evaluations

All the participants completed an evaluation form that used a Likert scale (range: 1-5) to evaluate the “ACE” process and had the opportunity for free-text comments on the form.

The responses were generally positive with the students deeming the “ACE” format as comprehensive, with a median score 4 (IQR: 4–4). The students reported that it identified their strengths and weaknesses, with a median of 4 (IQR: 4.5–5). It was generally agreed that the “ACE” was superior to the OSCE in assessing their management of an entire patient episode format, with a median of 4 (IQR: 4.5–5). Students reported that the “ACE” prepared them for the intern role, with a median of 4 (IQR: 4–5). Finally, students found the simulated patients/actors realistic and credible, with a median score of 4 (IQR: 4–4). This alternative integrative approach to assessing clinical skills “ACE” had a marked effect on participant’s self-reported confidence in their ability to undertake the intern role. Participants reported very positive attitudes to this new method of assessment, specifically relating to the integration of previously acquired knowledge and skills. The participants found the unfamiliar actors very authentic but felt the lack of clinical signs was a limitation. In summary, participants were very positive to this new assessment approach stating that it “drew all the information together” and highlighted their need to manage an entire patient episode “which was a scary albeit real concept.”

## DISCUSSION

Assessments should help focus learning and identify students’ strengths and weakness over the entire program with the offer of remedial teaching where indicated. This approach will produce medical graduates who apply their practical skills in such a manner that the public will have confidence in them. Future assessment methods will need to be reflective of a student’s actual abilities and indicate that the student is safe to practice on graduating.

Reliability is an essential component of the validity of any assessment. The “ACE” format proved a reliable assessment achieving an IRR Cronbach  $\alpha$  of 0.907 (CI: 0.766–1). The high IRR might be a consequence of all the examiners being surgeons. This finding is a further improvement on Luiz’s reported IRR of 0.89.<sup>14</sup>

The “ACE” format presented students with an acute surgical case (acute abdominal pain), whereas the long case format presents students with a surgical patient after an acute episode with postoperative complications. Thus, these examination methods may complement each other measuring new and long standing surgical conditions as required in the health care setting.

The “ACE” format failed a high number of students. Overall, 13 students were failed on the grounds of not achieving a score of 3/5 and were deemed at a level less than

required for intern/foundation year 1 practice by 2 independent examiners. The addition of a patient safety measure by avoiding an egregious error as pass criteria resulted in a further 14 students failing. There was strong statistical support to suggest that the distribution of gross “ACE” grades was different between the students who did and did not consider patient safety measures. This suggests that there is value in including the requirement to consider patient safety and avoid an egregious error in the determination of the final grade. This is not currently an assessment criterion in the long case examination or OSCE. It is suggested that the inclusion of the avoidance of an egregious error when compiling differential diagnoses in the examination standard setting program may increase specificity—this is an area that warrants further investigation.

There are 3 possible explanations for the observed odds ratio (33:1) of student failures between the “ACE” and the long case formats. The first explanation is that the “ACE” format was too difficult for the students. The second explanation is that the long case format was too easy.

The third explanation is that the “ACE” format significantly prepared students for the long case format.

The authors do not believe that the first explanation is valid owing to the stringent standard setting in the “ACE” format. The “ACE” timing and format may have unwittingly remediated students, thus acting as a teaching aid. Students who failed the “ACE” may have benefited from the comprehensive feedback and so were able to target their revision with a potential to markedly improve their subsequent performance in the long case examination. This may explain the reduced number of failures in the long case format.

Ensuring specificity in clinical skills examination in medicine remains a challenge to educators. The results of this study indicate that students and examiners found the “ACE” format comprehensive and discriminatory by identifying those not yet at a level expected to provide safe patient care. This examination format also has potential as a practice/mock examination to remediate those performing below the intern/foundation year 1 level, facilitating all the graduates of medical degrees to meet regulatory standards for entry into the workforce.

## LIMITATIONS

The researchers and examination coordinators (M.M. and A.E.G.) have a close working relationship with the participants, which may have influenced their decision to participate. Having the opportunity for formative assessment with written and verbal feedback may explain the participant’s positivity toward the examination rather than the “ACE” format itself. The participants only experienced 1 of 2 cases, so extrapolation to overall ability is limited. More cases are needed to examine the potential

generalizability of performance. Robust comparisons between the “ACE” and traditional long case examination could not be made because of the time difference between examinations (4 mo), and simulated patients were used in the “ACE” format for standardization, whereas the long case format uses real patients for authenticity.

## CONCLUSION

Ensuring that medical graduates are safe to practice it is mandatory for all training institutions. It is essential that the final clinical skills examination has good sensitivity and more importantly specificity to ensure that students who pass are safe to practice. The current “gold standard” certification examination in medicine (traditional long case format) has despite many adjustments, well-documented limitations.<sup>7-10</sup> The authenticity of assessing a real patient in a contextually relevant environment of an acute clinical setting is unsurpassable. However, medical training class sizes have increased in recent years to meet societal requirements for sufficient doctors to manage an ageing

population.<sup>18</sup> This places ever increasing demands on finite resources. Access to sufficient patients and availability of sufficient subject experts may restrict the potential of the long case examination methodology going forward. The “ACE” is a feasible adjunctive assessment format at the end of a primary degree.<sup>15</sup> It involves an integrative process and expands the domains evaluated in current assessment methods. Going forward, more cases are required to sample the curriculum more widely. Ultimately, students need to be assessed on multiple cases, on multiple occasions using multiple assessors.<sup>19</sup>

## ACKNOWLEDGMENTS

Many thanks to Dr Michael Bennett who worked on the original “ACE” feasibility examination; to the School of Medicine for the Education Grant; to Dr Helen McCrae for the “PAME” framework, which underpinned this study; to all the students for participating; to all the staff who gave their time to examine; and to all the invigilators who assisted in the running of the “ACE.”

## APPENDIX 1. “ACE” MARKING SHEETS BASED ON PAME

Examiner ID #:	STICKER/ID
----------------	------------

### Examiner Global Rating Scale—Encounter 1: History and Examination

#### HISTORY

1 unacceptable <b>FAIL</b>	2 inadequate <b>FAIL</b>	3 barely adequate for intern/FY1 practice <b>PASS</b>	4 adequate <b>2H</b>	5 outstanding <b>1H</b>
----------------------------------	--------------------------------	---	----------------------------	-------------------------------

#### PHYSICAL

1 unacceptable <b>FAIL</b>	2 inadequate <b>FAIL</b>	3 barely adequate for intern/FY1 practice <b>PASS</b>	4 adequate <b>2H</b>	5 outstanding <b>1H</b>
----------------------------------	--------------------------------	---	----------------------------	-------------------------------

#### KNOWLEDGE

1 unacceptable <b>FAIL</b>	2 inadequate <b>FAIL</b>	3 barely adequate for intern/FY1 practice <b>PASS</b>	4 adequate <b>2H</b>	5 outstanding <b>1H</b>
----------------------------------	--------------------------------	---	----------------------------	-------------------------------

#### COMMUNICATION SKILLS

1 unacceptable <b>FAIL</b>	2 inadequate <b>FAIL</b>	3 barely adequate for intern/FY1 practice <b>PASS</b>	4 adequate <b>2H</b>	5 outstanding <b>1H</b>
----------------------------------	--------------------------------	---	----------------------------	-------------------------------

#### TIME MANAGEMENT

1 unacceptable <b>FAIL</b>	2 inadequate <b>FAIL</b>	3 barely adequate for intern/FY1 practice <b>PASS</b>	4 adequate <b>2H</b>	5 outstanding <b>1H</b>
----------------------------------	--------------------------------	---	----------------------------	-------------------------------

#### COMMENTS

(Please make comments on reverse of sheet on both good and bad aspects of the students' performance):



**Examiner Global Rating Scale—Encounter 2: Investigations**USE OF RESOURCES

1 unacceptable <b>FAIL</b>	2 inadequate <b>FAIL</b>	3 barely adequate for intern/FY1 practice <b>PASS</b>	4 adequate <b>2H</b>	5 outstanding <b>1H</b>
----------------------------------	--------------------------------	---	----------------------------	-------------------------------

INTERPRETATION

1 unacceptable <b>FAIL</b>	2 inadequate <b>FAIL</b>	3 barely adequate for intern/FY1 practice <b>PASS</b>	4 adequate <b>2H</b>	5 outstanding <b>1H</b>
----------------------------------	--------------------------------	---	----------------------------	-------------------------------

**COMMENTS**

(Please make comments on both good and bad aspects of student's performance):

**Examiner Rating Scale—Encounter 3: Oral Examination Diagnosis and Management**Knowledge

1 unacceptable <b>FAIL</b>	2 inadequate <b>FAIL</b>	3 barely adequate for intern/FY1 practice <b>PASS</b>	4 adequate <b>2H</b>	5 outstanding <b>1H</b>
----------------------------------	--------------------------------	---	----------------------------	-------------------------------

Judgment

1 unacceptable <b>FAIL</b>	2 inadequate <b>FAIL</b>	3 barely adequate for intern/FY1 practice <b>PASS</b>	4 adequate <b>2H</b>	5 outstanding <b>1H</b>
----------------------------------	--------------------------------	---	----------------------------	-------------------------------

Communication with Examiner

1 unacceptable <b>FAIL</b>	2 inadequate <b>FAIL</b>	3 barely adequate for intern/FY1 practice <b>PASS</b>	4 adequate <b>2H</b>	5 outstanding <b>1H</b>
----------------------------------	--------------------------------	---	----------------------------	-------------------------------

**COMMENTS**

(Please make comments on both good and bad aspects of student's performance):

**Examiner Rating Scale—Encounter 4: Review and Management**MANAGEMENT PLAN

1 unacceptable <b>FAIL</b>	2 inadequate <b>FAIL</b>	3 barely adequate for intern/FY1 practice <b>PASS</b>	4 adequate <b>2H</b>	5 outstanding <b>1H</b>
----------------------------------	--------------------------------	---	----------------------------	-------------------------------

KNOWLEDGE

1 unacceptable <b>FAIL</b>	2 inadequate <b>FAIL</b>	3 barely adequate for intern/FY1 practice <b>PASS</b>	4 adequate <b>2H</b>	5 outstanding <b>1H</b>
----------------------------------	--------------------------------	---	----------------------------	-------------------------------

COMMUNICATION SKILLS

1 unacceptable <b>FAIL</b>	2 inadequate <b>FAIL</b>	3 barely adequate for intern/FY1 practice <b>PASS</b>	4 adequate <b>2H</b>	5 outstanding <b>1H</b>
----------------------------------	--------------------------------	---	----------------------------	-------------------------------

RESPONSE TO PATIENT CONCERNS/WISHES

1 unacceptable <b>FAIL</b>	2 inadequate <b>FAIL</b>	3 barely adequate for intern/FY1 practice <b>PASS</b>	4 adequate <b>2H</b>	5 outstanding <b>1H</b>
----------------------------------	--------------------------------	---	----------------------------	-------------------------------

**COMMENTS**

(Please make comments on both good and bad aspects of students performance):



7. Van der Vleuten C. Validity of final examinations in undergraduate medical training. *Br Med J*. 2000;321(7270):1217-1293.
8. Eva KW. On the generality of specificity. *Med Educ*. 2003;37(7):587-588.
9. Wass V, van der Vleuten C. The long case. *Med Educ*. 2004;38(11):1176-1180.
10. Wass V, Jolly B. Research paper: does observation add to the validity of the long case? *Med Educ*. 2001;35(8):729-734.
11. Norcini J. The validity of the long case. *Med Educ*. 2001;35(8):720-721.
12. Gleeson F. The effect of immediate feedback on clinical skills using the OSLER. In: Rothman AI, Cohen R, eds. *Proceedings of the Sixth Ottawa Conference of Medical Education*. Toronto: University of Toronto Bookshop Publishing; 1994. p. P412-P415.
13. Hamdy H, Prasad K, Williams R, Salih FA. Reliability and validity of the direct observation clinical encounter examination (DOCEE). *Med Educ*. 2003;37(3):205-212.
14. Luiz EAT, Roberto OD, Fernando CF, et al. A standardised structured long case examination of clinical competence of senior medical students. *Med Teach*. 2000;22(4):380-385.
15. Morris M, Bennett M, Hennessy M, Conlon KC, Ridgway PF. An alternative certification examination "ACE": can post graduate methods be used to assess clinical skills in medical under graduates? *Br J Med Med Res*. 2013;3(4):1806-1817.
16. MacRae HM, Cohen R, Regehr G, Reznick R, Burstein M. A new assessment tool: the patient assessment and management examination. *Surgery*. 1997;122(2):335-344.
17. Greene J. *Learning to Use Statistical Tests in Psychology*. 2nd ed Bucks, UK: Open University Press; 1999.
18. Fottrell Report on Undergraduate Teaching, Medical Education in Ireland: A New Direction. 2006.
19. Van der Vleuten CPM. Assessing professional competence: from methods to programmes. *Lancet*. 2005;39(9):309-317.