ORIGINAL ARTICLE

Final voluntary assessment for Traumatology and Orthopaedic Surgery medical residents: A report on the results and a look at the future

A. Urda a, L. Ezquerra b, J. Albareda b, J. Baeza-Noci c, A. Blanco d, E. Cáceres e, M. Martínez-Grande f, J. Nardi e, A. Yunta g, F. Marco a, *

a Servicio de Cirugía Ortopédica y Traumatología, Hospital Clínico San Carlos, Madrid, Spain
b Servicio de Cirugía Ortopédica y Traumatología, Hospital Clínico Universitario Lozano Blesa, Zaragoza, Spain
c Servicio de Cirugía Ortopédica y Traumatología, Hospital Virgen del Consuelo, Valencia, Spain
d Servicio de Cirugía Ortopédica y Traumatología, Complejo Asistencial Universitario de Burgos, Burgos, Spain
e Departamento de Cirugía Ortopédica y Traumatología, Ciudad Sanitaria Vall D’Hebron, Barcelona, Spain
f Servicio de Cirugía Ortopédica y Traumatología, Clínica Ubarmin, Pamplona, Navarra, Spain
g Servicio de Cirugía Ortopédica y Traumatología, Fundación Hospital Asilo de Granollers, Granollers, Barcelona, Spain

Received 18 December 2011; accepted 19 December 2011

KEYWORDS
MIR; Specialist training; Residency programme; Final exam; Orthopaedic Surgery and Traumatology

Abstract

Introduction: The idea of establishing an examination that accredits the training of the specialists in orthopaedic surgery at the end of their educational period as residents is subject to controversy. With the aim of encouraging the development of this examination, the present members of the National Commission of the Specialty of Orthopaedics (CNE) have reviewed the results obtained in previous examination.

Materials and methods: The results of the voluntary final exam for Orthopaedics participants, and of the surveys of participants opinions for the years 2006–2011 are presented.

Results: The total number of participants was 231, growing from 19 in 2005 to 71 in 2011. The overall mean score in the period reviewed (2006–2011) was 6.72 out of 10 points. In these 6 years, 9 participants failed (4.25%). The mean score for the test was 7.57. The overall mean score of the oral exam was 6.57. The worst results were obtained in the general knowledge section of the oral examination. Nobody has ever failed the section on reconstruction of the lower extremity. The upper extremity section in the oral examination achieves the best average results.


* Corresponding author.
E-mail address: fmarco@med.ucm.es (F. Marco).

1988-8856/$ - see front matter © 2011 SECOT. Published by Elsevier España, S.L. All rights reserved.
Discussion: The examination has established its place in the structure of orthopaedic resident training in our country, even without making it official by the Administration. We must positively acknowledge those candidates who chose to take the examination and be evaluated by their peers in order to prove their qualification and distinction. This recognition will be the best encouragement for future generations.

© 2011 SECOT. Published by Elsevier España, S.L. All rights reserved.

PALABRAS CLAVE
MIR;
Formación especializada;
Residencia;
Examen final;
Cirugía Ortopédica y Traumatología

Evaluación final voluntaria para médicos residentes de Cirugía Ortopédica y Traumatología: balance de resultados y mirada al futuro

Resumen
Introducción: La conveniencia de establecer una prueba que acredite la formación de los especialistas en Cirugía Ortopédica y Traumatología (COT) al término de su periodo de formación como médicos internos residentes es objeto de controversia. Con la intención de favorecer y estimular el desarrollo de esta prueba, los actuales vocales de la Comisión Nacional de la Especialidad (CNE) han realizado una revisión de los datos obtenidos en las pasadas ediciones.

Material y métodos: Resultados de las pruebas de evaluación voluntaria de final de residencia y de las encuestas de valoración rellenadas por los participantes de la prueba entre los años 2006 y 2011.

Resultados: En total han intervenido 231 personas en la evaluación, incrementándose el número de participantes desde 19 en 2005, hasta 71 en 2011. La media de puntuación global a lo largo del periodo analizado (2006-2011) es de 6,72 puntos. Los no aptos en estos 6 años han sido 9 participantes (4,25%). La media de puntuación en el test es de 7,57 puntos. La puntuación global media de las mesas es de 6,57. Generalidades es la mesa que obtiene peores resultados. La mesa de miembro inferior reconstructivo no ha sido suspendida nunca. La mesa de miembro superior es la que obtiene los mejores resultados medios.

Discusión: La prueba está arraigada en la estructura de la formación especializada de la COT en nuestro país. Debemos reconocer favorablemente a aquellos candidatos que optan por presentarse. Este reconocimiento será el mejor estímulo para las futuras promociones.

© 2011 SECOT. Publicado por Elsevier España, S.L. Todos los derechos reservados.

Introduction

Whether it is a good idea to establish a test that certifies the training of Orthopaedic Surgery and Traumatology (OST) specialists at the end of their training period as medical residents (médicos internos residentes, MIR, in Spanish) has been and remains controversial. Numerous colleagues have weighed in as detractors or fans of such a test. Among the later, the late Professor Luis Munuera was the one who promoted what is known as the "Voluntary Final MIR Assessment in Orthopaedic Surgery and Traumatology", which had its 7th session in 2011.

When the MIR system was set up, Royal Decree 127/1984 established the possibility of substituting the final examination by continuous annual assessments. However, the option of the test gained more and more ground until it was expressed in the latest legal update on specialised training in 2008. Royal Decree 183/2008 indicated: "When the final assessment of the junior doctor period is positive, the national commission of the corresponding specialty will give a test to the junior doctors applying for it, to obtain one of the following marks, in order of increasing qualification: a) Distinguished with mention by the national commission for the specialty, and b) Distinguished with Special Mention by the national commission for the specialty".

The National Commission for the Specialty (NCS) has been a pioneer among the rest of the specialties in carrying out a "running-in" phase that will allow it to face the official proposals with sufficient experience. It is also true that the whole-hearted support of its scientific society, SECOT (Spanish Society of Orthopaedic Surgery and Traumatology), which has made all the material and human resources needed available, has facilitated its survival. Without this generous effort, this would not have been possible. The dissemination of this experience has been spreading annually through the partial transmission of the results to the Academic Units in each centre. However, various voices – and some as accredited as that of Guerado, in an editorial in our journal3 – have suggested that it is necessary to collect an analysis of what has happened until now. In the attempt to favour and stimulate the development of a test in which the passage of time has led us to believe, the current NCS representatives have gathered and revised the data.

Material and methods

The NCS members collected the data presented in this article throughout the 2005–2011 period. The quality of the data has varied, becoming more complex over time; consequently, some information may apply to only the final segments of this period. Going even farther than just the
test results, forms and questionnaires filled out by the participants at the time of assessment have been used. The Candidate anonymity was maintained throughout the questionnaire process.

The first session of the assessment test at the end of the residency was in 2005, during the national SECOT congress in Seville. At that time, no statistics were gathered on the 19 participants. Consequently, the data analysed henceforth correspond to the following 6 annual sessions, from 2006 to 2011.

The test consists of 2 parts. The first corresponds to a multiple-choice test, with only 1 valid answer out of 5 possible ones. There are 50 questions and the candidates have 1 h to finish. There are no negative points for incorrect responses. The second part is an oral examination divided into 6 specific assessment “tables” or sections: general knowledge, upper limb, lower limb reconstruction, lower limb trauma, spinal column and paediatric orthopaedics. The candidates respond to questions on practical hypothetical situations posed by 2 assessors per section. The time allotted for each oral exam section is 20 min and the score is the sum of the various aspects assessed (initiative and autonomy, ability to analyse, rational use of complementary tests, ability to communicate, etc.) in addition to the assessment of knowledge. Individual results are not made public and only the candidate receives the certification.

Results

In the 7 sessions held to date, 231 individuals have participated, of which the sample analysed corresponds to the 212 candidates from the 2006 to 2011 period. The number of participants has grown systematically from 19 in 2005; the 2011 session was the most numerous, with 71 MIR, which represents an increase of 209% compared to 2010, in which there were 34 candidates (Fig. 1).

The participants have been mostly males throughout all the editions, with a mean 72%. However, the percentage of females has gradually been increasing, reaching 42% in the 2011 session.

Data on the hospitals from which the candidates came began to be collected in 2007, categorising the hospitals based on the number of residents per year that they train. Two hospital categories were established: hospitals that admit 3 or more residents a year, and hospitals that admit 1 or 2. In 2007, most of the test candidates came from hospitals training 3 or more residents per year. This changed in 2008 and has remained constant throughout the following sessions; since then, a mean of 69% of the participants have come from hospitals with 1 or 2 residents a year since then.

Participant geographical origin began to be collected in 2008 and has increased over the sessions. In 2008, there were candidates from 8 autonomous communities in Spain, with Cataluña being the community that supplied the greatest number of MIR, with 46%. In 2009 and 2010, the junior doctors came from 9 autonomous communities and, in 2011, this figure rose to 13. Throughout the sessions, Cataluña,
Madrid and the Valencia Community have been the ones (in that order) that sent the most candidates (Fig. 2).

The overall test results for the 6 sessions take into account both the test and oral scores, with the final result being the average of the 7 marks. The final results are classified, in accordance with Royal Decree 183/20082 (Articles 25.2 and 26.2), in Fail if the score is less than 5, Pass between 5.0 and 6.9, Mention between 7.0 and 7.9 and Special Mention for the scores between 8.0 and 10 (Fig. 3).

The mean overall score for all the sessions is 6.72 points. The best year was 2006, with a mean of 7.10, and the worst, 2008 with 6.42 points. In 2006, the candidates achieved the highest number of Special Mentions, with 33% achieving this distinction. However, in 2009 none of the candidates managed to reach a mean mark above 8 to receive this distinction. Achieving Special Mention seems to be more difficult as the sessions go by, given that from 33% who received it in 2006, the percentage has gradually been decreasing: 24% in 2007, 6.9% in 2008, 0% in 2009 and 2.94% in 2010, up to 2011, when 2.82% of the young specialists obtained this maximum qualification (Fig. 4).

The distribution of the results over the years is shown in Fig. 4. Note how, in 2009, there was no candidate who received a Fail mark, but there was none who achieved a Special Mention either.

The candidates with Fail in these 6 years were 9, which represents 4.25% of the total 212 individuals taking the test whose data were analysed. The year with the greatest number of Fail marks was 2008, with 6.9% of the MIR. In 2009, no candidate received a Fail.

Looking individually at each of the sections that make up the exam, the mean score on the test (the multiple-choice part) over the 6 sessions was 7.57 points (Fig. 5). In this section, the participants obtained consistently good results, as evidenced by the fact that only 2 MIR from the 212 total failed the test, 1 in 2007 and another in 2009. The year with the best scores was 2006, with a mean of 8.05; the worst was 2009, with a mean of 6.97 points.

Fig. 5 shows the mean scores of the tests, of the oral assessments and of the overall results for the sessions from 2006 to 2011.

The mean overall score from the oral examinations was 6.57, ranging from 6.17 in 2007 to 7.10 in 2009 (Fig. 5). The section in which the candidates obtained the best mean result was that of the upper limb, with a score of 7.14; followed by lower limb traumatisms, with 6.89; lower limb reconstruction, 6.75; spinal column, 6.67; paediatric orthopaedics, 6.22; and general knowledge, 5.72.

The “General Knowledge” oral exam is the one that, since 2006, obtained the worst results. The mean score for the 6 sessions was 5.72 points. The year with the best result was 2006, with 6.40, while the mean score was only 5.25 points in 2011. It is also the oral section at which the largest number of participants failed over all the sessions, with 58 of the 212 candidates, indicating that 27.36% of the total failed in this section. This percentage ranged from 0 in 2006, when nobody failed the table, up to 37.93% in 2008. Since that time, it has been the segment with the worst mean score every year. In 2006, as stated above, no resident failed the examination, although at the spinal section they performed worst (Fig. 6).

In contrast to the General Knowledge oral examination stands the “Lower Limb Reconstruction” section, in which no participant has failed during the 6 sessions. It is third in mean score, with 6.75 points. The lowest mean score was obtained in 2007, with 5.50 points, and the best, in 2009, with 7.78.

The “Lower Limb Trauma” oral assessment section has a similar 6-session mean score, with 6.89 points, behind only the upper limb section. The performance of the MIR is fairly satisfactory here; in fact, participants failed on only 2 occasions, in 2007 and 2008, with 4% and 20.69%, respectively. In the rest of the sessions, no candidate has ever failed this lower limb section. In 2009, the mean score was 8.10 points, which is the best mean score obtained to date among all the sections from all the years.
The "Upper Limb" assessment section obtained the best results in the 6 sessions, with a score of 7.14. The best year was 2006, with 7.74 points; the worst, 2008, with 6.60 points. There were few Fail marks given in this oral section either. Candidates failed on only 2 occasions, in 2008 and 2011, in which 10.34% and 9.86% of the doctors failed this assessment, respectively.

The "Paediatric Orthopaedics" oral exam section is the one with the second worse performance, after the general knowledge oral, with a mean of 6.22 points. In 2007, a failing mean score was obtained in this oral, with 4.00 points. It was the only case in which a section was failed in an overall manner during the entire 6 sessions. It is also the second oral section, after general knowledge, that the largest number of MIR failed over the study period, with 29 of the 212 candidates (13.68%) not passing. At least 1 candidate failed this oral section (as happened with the spinal column oral) every year. The year with the highest percentage of Fail marks (24%) was 2007 and that with the lowest (4.76%) in 2006.

In the "Spinal Column" oral assessment section, intermediate results were obtained, behind the upper limb and the 2 lower limb sections, but ahead of the general and paediatric orthopaedics sections. The mean score was 6.67 points. The score obtained was fairly homogeneous (approximately 6.5 points) over the sessions, except for 2007, in which the second highest mean score obtained in all sessions in an oral assessment exam section was obtained, with 8.00. As happened with paediatric orthopaedics, at least 1 candidate failed this section in all the sessions. The worst year was 2006, when 28.57% of the MIR received a failing mark, which made it the oral exam section with the largest number of Fail marks that year. The best year was 2009, in which only 1 (3.13%) candidate failed.

Returning to the overall oral exam data, with respect to Special Mention (Fig. 7), lower limb trauma was the section in which the most were achieved, 37.26% of the candidates. In descending order, next was the lower limb reconstruction section (28.30%), then upper limb (26.89%), spinal column (23.58%) and paediatric orthopaedics (23.11%) sections. The oral section in which the fewest Special Mention marks were obtained was general knowledge, with only 13.21% of the participants during these 6 years.

---

**Figure 6**  (A) Mean general knowledge oral examination score in the 6 sessions. (B) Percentage of Fail and Special Mentions in that oral exam section.

**Figure 7**  Percentage of Special Mentions in the different oral examination sections in the 2006-2011 sessions.
The percentage of Fails followed the inverse order (Fig. 8); the general knowledge oral was the section which the most junior doctors failed (27.36%), followed by paediatric orthopaedics (13.68%), spinal column (11.32%), upper limb (4.72%) and lower limb trauma (3.30%) sections. The best oral section was the already-mentioned case of lower limb reconstruction, at which no participant failed over the course of the 6 study sessions.

As additional information, we would like to mention the study material that the candidates indicated that they used as reference material to prepare for the test (data collected since 2009) (Table 1). Beginning in 2007, the MIR were also asked to indicate voluntarily information about their job prospects after obtaining the specialist certification. This information has now been collected since 2009, asking the candidates to specify the type of work contract they had (Tables 2 and 3). The candidates indicated that their main reason for taking the examination was their interest in self-assessing the knowledge achieved, with this motivation reaching 85% in the 2011 session. Finally, their overall opinion of the test (on a scale from 1 to 10) was 8.97 (2009), 8.29 (2010) and 8.7 (2011).

Discussion

Without presenting an exhaustive analysis of the data, there are various observations that we can make. The first is that the test is deeply entrenched in the structure of OST specialist training in Spain, even before the start of official examinations by the Administration. It has not only withstood the test of time, the number of candidates has increased (spectacularly so in the last year). If we take the positions offered in 2005 and those finishing the MIR
in 2011 (180 positions), the 71 participants in the last session mean that we assessed 40% of them. What is the reason for this? As always, it is due to various factors. In the first place, it reflects the success of the collaboration between 2 organisms, NCS and SECOT, which have put a project in which they believed before any personal interests. Within this common undertaking, the fact that SECOT obtains financial resources with the participation of the industry is always to be emphasised. The collaboration of the OST Academic Units, represented by their Heads of Service and Tutors, has been fundamental in encouraging the new specialists to take the test. Finally, the personal experience of those who have gone through the assessment, with a pleasant atmosphere for both parts, has made word-of-mouth a tool for progress. The participant questionnaire about their overall opinion of the test gave a mean of 8.65 out of a maximum positive score of 10 points.

Insofar as the participant profile, the gender segmentation reflects the sociodemographic characteristics in our specialty. The 2011 SECOT White Book indicated 83% males in OST and, in the assessment test, 72%. However, females have entered strongly into the OST, with 42% of the candidates in the last session. For assessment development, the participation of the Academic Units of the centres that train 1–2 residents a year is very important. The largest hospitals have gradually been losing proportion (27% in 2011, against 64% in 2007), perhaps indicating more dynamic training in smaller centres. With respect to the geographic origin of the participants, we can feel satisfied with the increase in the number of autonomous communities, which reached 13 in the last session. In contrast, the lack of candidates from Extremadura, Asturias, Rioja and Cantabria in all the sessions for which we have geographic data is notable. Considering percentages, we can compare the ratio of specialists trained in each community with the weight of their participation in the assessment test. In the 2011 MIR test, Cataluña offered 18% of the OST positions, but the candidates from this community account for 27% of the total throughout these 6 years. That is, Cataluña is very well represented, as compared to (for example) Andalucía, whose participation was only 6% of the total OST positions available, although its training offer was 19% of the total OST positions. To close out this observation, Madrid can be considered in perfect balance, with 16% of the overall participants and an offer of MIR positions of 16% in 2011.

The results that the candidates obtained follow a Gaussian bell curve distribution. Over the study period, the percent of MIR who received a Fail mark remained in the lower registers. Even so, it should be remembered that, given that this is a voluntary test, adequate preparation can be assumed. The test results have also become more selective in the qualifications of excellence and the percentage of Special Mention (equivalent to an ‘A’ or ‘1st class’) has gradually dropped to less than 3%. This point emphasises the work of the assessors that have taken part over the years in the process, with nearly 100 participations. All of them attend preparatory encounters prior to the assessment test and conclusion round-up sessions after it, which have permitted homogenising the oral assessments, to a degree. The observations extracted from these encounters have let them and their successors gradually perfect to the oral examination, maintaining the requisite cordial atmosphere with the candidates while sharpening their judgement as to the candidates’ capabilities. The assessors’ dedication in donating their time and knowledge selflessly deserves a grateful acknowledgement, which the NCS and SECOT (and, by extension, all of the colleagues) should give them (Fig. 9).

With respect to the specific results in each section of the test, there are also a few interesting observations. The NCS members prepared the test to evaluate basic theoretical knowledge and the consistency in candidate scores gives proof of this. The MIR passed the written questions with “B” marks consistently and a failing mark was seldom received. Turning to the mean in oral exam assessments, the results tended to be lower than those of the multiple-choice test. The failing order of the training shows that our young

Table 3 Type of contract held by the participants who were working. In 2010, up to 24% of the participants were working without a contract when they took the test.

<table>
<thead>
<tr>
<th></th>
<th>Standard (full-time)</th>
<th>Locum</th>
<th>Temporary</th>
<th>No contract</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>19%</td>
<td>31%</td>
<td>38%</td>
<td>3%</td>
<td>9%</td>
</tr>
<tr>
<td>2010</td>
<td>15%</td>
<td>15%</td>
<td>43%</td>
<td>24%</td>
<td>3%</td>
</tr>
<tr>
<td>2011</td>
<td>12%</td>
<td>16%</td>
<td>51%</td>
<td>6%</td>
<td>15%</td>
</tr>
</tbody>
</table>

Figure 9 Partial view of the oral examination hall (2011) with 3 of the assessment section tables.
specialists were capable of resolving limb problems better than those of the spinal column, and were even less proficient with cases of paediatric orthopaedics. Finally, they had the greatest difficulty in general knowledge of scientific concepts of our discipline (which also include tumoural and infectious pathology).

This oral exam result scaling should serve as an indication for the OST Academic Units as to the training we offer. Although these are select candidates, as their overall marks indicate, the imbalance in the different areas of our specialty should gradually be corrected. It is difficult to agree completely with Epeldegui, who, motivated by his love of paediatric orthopaedics, incorrectly combines the development of the core training in the MIR system with the possible areas of specific training and even the idea of defining non-specific rotations, segmenting the periods of specific training dedicated completely to the various aspects of our specialty. Even circumscribing this author’s right to give an opinion due to not being specifically dedicated to the activity, the NCS needs to recommend that the individuals responsible for each Academic Unit provide appropriate supervision for the acquiring of knowledge in these 3 fields (spinal column, general knowledge and paediatric orthopaedics) that are left slightly behind by the others. Our training programme, published in 2007, clearly specifies that “the objectives need to be oriented in their content towards the training of a basic specialist”. González, in his work on training in spine surgery in Spain, acknowledges the necessary difference between the MIR training period (using elective rotations to round out some specific aspect) and the later development of the specialist in the area they choose, using the future Academic Units accredited for specific training to do so.

We would like to point out, as complementary information gained from the various questionnaires given to the candidates, how popular the “SECOT Manual on Orthopaedic Surgery and Traumatology” was. It was cited as the preferred reference for preparing the test. The relevance of the text should constitute a reward for our Society's significant effort in preparing it. Also noteworthy was Delgado’s text, “Orthopaedic Surgery and Traumatology”. As to the job opportunities of our recently graduated doctors, a greater difficulty in joining the labour market can be seen in the last years. This should serve as a warning to all the organisms involved about the need to adjust training, as much as possible, to the true demand for OST specialists.

Finally, it is essential for us to speculate on the future of the test. This will probably require turning our attention to 3 aspects. First, there is the option of the health administration in Spain. As we have seen, the test has been entered officially into legal regulations. Its development, however, still needs to be specified and there are many pending tasks that can affect it, such as the already mentioned “Royal Decree Project regulating the incorporation of core areas in the training of specific specialties in the Health Sciences, core re-specialisation and specific training areas”. If this decree prospers, it would make it obligatory to at least prepare new programmes for the specialties; this would probably delay the implementation of all the pending aspects contemplated in Royal Decree 183/20082, including this assessment. In second place, however, NCS and SECOT still have to recognise the MIR test officially. Both bodies have shown their firm support of it and its current success should spur them to continue to support and improve it in future sessions. In third place, but probably as important as the others, there is our integration into the guidelines set by the European Union. As Cordero pointed out, the curriculum is imbalanced among the different member states, even as to the length of specialised training. However, there is great agreement as to an obligatory final exam, with that author listing 15 countries that perform it and indicating that it is voluntary in Spain and Sweden. Consequently, we should not imagine that such assessment is going to disappear, but that it is going to be consolidated. What is more, from this European point of view, the European Board of Orthopaedics and Traumatology (EBOT) Exam turned 10 years old in 2010. Its structure imitates our assessment and this is not by chance, but rather thanks to NCS and SECOT’s vision of the future. Integrating the tests for OST specialists in the countries of the European Union is an objective that should make us all excited. With all of this, it falls to our professional community to assert the value of the test. We need to give favourable recognition to the candidates who choose to present themselves before their peers to be assessed at the end of their education and demonstrate their training and excellence. This recognition would be the best incentive for future classes.

Level of evidence
Level of evidence IV.

Ethical Disclosures
Protection of human and animal subjects. The authors declare that no experiments were performed on humans or animals for this investigation.
Confidentiality of Data. The authors declare that no patient data appears in this article.
Right to privacy and informed consent. The authors declare that no patient data appears in this article.

Conflict of interest
The authors declare that they have no conflicts of interest.

References
2. Real Decreto 183/2008, de 8 de febrero, por el que se determinan y clasifican las especialidades en Ciencias de la Salud y se desarrollan determinados aspectos del sistema de formación sanitaria especializada. Boletín Oficial del Estado, número 45, 21/2/2008.
4. SECOT. Libro Blanco SECOT de La Cirugía Ortopédica y Traumatología en España; 2011.