



Memoria para optar al Grado de Doctor:

Comparación de la eficacia del aprendizaje basado en casos/problema frente al método tradicional de lección magistral para la enseñanza de la Anestesiología en pregrado, postgrado y Formación Médica Continuada.

Doctorando: Enrique J Carrero Cardenal

Directora: Profesora Carmen Gomar Sancho

Departament de Cirurgia i Especialitats Quirúrgiques

Facultad de Medicina. Universitat de Barcelona

Barcelona, junio 2009



UNIVERSITAT DE BARCELONA



Facultat de Medicina
Dept. Cirurgia i Especialitats Quirúrgiques
Casanova, 143
08036 Barcelona
Tf. 934039684

La Dra **Carmen Gomar Sancho**, como Directora del trabajo de Tesis doctoral:
Comparación de la eficacia del aprendizaje basado en casos/problema frente al método tradicional de lección magistral para la enseñanza de la Anestesiología en pregrado, postgrado y Formación Médica Continuada.

Certifica:

Que esta Tesis ha sido realizada por **Enrique J Carrero Cardenal**, bajo mi dirección, para optar al Grado de Doctor.

Los artículos que se incluyen en esta tesis cumplen las condiciones vigentes en la Facultad de Medicina de la Universitat de Barcelona para la presentación de tesis doctorales por publicaciones.

En Barcelona, a.....2009.

Firma

“Para tener la certeza absoluta de algo, uno debe saberlo todo o nada de ese algo” (Olin Miller, humorista)

“La clave de una buena investigación comienza por ser flexible y aceptar las limitaciones de cualquier metodología en cualquier campo” (Sybille K Lechner, 2001)

“La investigación en educación médica importa” (Hardem RM, 2000)

Presentación

Esta memoria presenta el diseño, la metodología y resultados de un estudio amplio de comparación de dos métodos de enseñanza de la Anestesiología: el método de enseñanza tradicional basado en la *clase magistral* y el *aprendizaje basado en la discusión de casos/problema*. Se ha estructurado en formato “por compendio de publicaciones”.

La introducción tiene por objeto enmarcar el estado actual del tema y definir los aspectos metodológicos del diseño randomizado para responder a las preguntas planteadas en los objetivos del estudio. En el apartado de material y métodos se describe la planificación y el diseño del estudio. Este apartado se ha estructurado en dos partes, una general y otra específica para cada artículo.

Se presentan tres artículos que se han elaborado a partir de la implementación del trabajo de campo y la explotación estadística de los datos obtenidos. El primer artículo compara los dos métodos en la enseñanza del *soporte vital básico*, uno de los contenidos más trascendentes en la formación troncal de Anestesiología y Reanimación en el contexto de la formación *pregrado*. El segundo artículo se enmarca en la formación *post-grado* de los residentes de Anestesiología de primer año; el tema sobre el que se investiga la eficacia de ambos métodos de enseñanza es la *valoración pre-anestésica*. El tercer artículo presenta los resultados del estudio en la *Formación Médica Continuada* de la Anestesiología. El contenido estudiado es el *embolismo aéreo* como paradigma de incidente crítico intraoperatorio.

Los resultados de los tres artículos se han analizado de forma conjunta en un apartado de discusión general. En este apartado se comentan los aspectos más importantes de la discusión concreta que se ha hecho y presentado para cada uno de los artículos. El objetivo es recapitular sobre los artículos de manera sistemática e interrelacionarlos para cumplir el objetivo general de esta tesis. Finalmente se exponen las conclusiones que se derivan de la investigación.

En el apartado de bibliografía general se incluyen las referencias utilizadas en la introducción y la discusión generales.

Como información complementaria se ha incorporado un apéndice con las aportaciones presentadas a reuniones científicas originadas por esta investigación.

Agradecimientos

A la **Profesora Carmen Gomar Sáncho**. Recuerdo que siendo yo estudiante de medicina ya me impresionó como demostraba *saber enseñar*... Desde entonces, y durante lo veinte años de profesión que he tenido el privilegio de compartir con ella, no ha habido un solo día que haya dejado de sorprenderme por su capacidad infinita de innovación, profesionalidad y dedicación. A la profesora Gomar le debo mucho de lo que soy como profesional y como docente, si no todo; sólo espero, algún día, ser capaz de transmitir a los demás el mismo espíritu que ella me ha enseñado.

A la **Profesora Wilma Penzo** le tengo que agradecer el gran entusiasmo y predisposición que siempre ha demostrado hacia este proyecto. Wilma: gracias por tus excelentes consejos y por enseñarnos lo importante que es la metodología en la investigación docente.

Quiero dar las gracias a los co-autores de los tres artículos por su inestimable ayuda. A la **Dra Fábregas** y al **Dr Valero**, por su ejemplo de trabajo profesional, por sus valiosos consejos y su colaboración y comprensión en todo momento; al **Dr Sanchez-Etayo** por su gran entusiasmo en el trabajo realizado; a la **Dra Rull** por su interés e infatigable capacidad de trabajo y a los **Dres Castillo y Villalonga** por todas las horas que hemos compartido enseñando y aprendiendo, por su espíritu crítico y por su amistad.

Este trabajo no hubiera sido posible sin la colaboración de los **estudiantes de medicina, los residentes de Anestesiología y los especialistas que aceptaron participar en los estudios**. A todos ellos, gracias.

Quiero agradecer al **Departament de Cirurgia i Especialitats Quirúrgiques de la Universitat de Barcelona**, a la **Societat Catalana d'Anestesiologia, Reanimació i Terapèutica del Dolor** y al **Comité para la Educación Europea en Anestesiología** las facilidades proporcionadas para el desarrollo de los tres estudios.

Son muchos los **compañeros de trabajo y amigos** que me han *sufrido* durante los años que ha durado este proyecto... Siempre he encontrado en ellos apoyo y comprensión y por eso se merecen mi aplauso.

A **mis padres**, gracias a ellos he aprendido que lo más importante en la vida no es lo que consigues sino sentirte bien contigo mismo mientras te esfuerzas por conseguirlo.

Raúl, quiero que sepas que tu ejemplo me ha acompañado muchos días de trabajo y que no he desfallecido gracias a ti.

A **Chero, Moko, Lola, Martina e Irina**.... por vuestras infinitas horas de compañía, por vuestro amor desinteresado e inagotable, por compartir vuestra vida con nosotros.

A todos los que se esfuerzan por mejorar; a todos los que les comprenden y estimulan. A los que ofrecen el esfuerzo de su investigación al servicio de los demás...

“Amor Eterno”

Podrá nublarse el sol eternamente;
Podrá secarse en un instante el mar;
Podrá romperse el eje de la tierra
Como un débil cristal.
¡todo sucederá! Podrá la muerte
Cubirme con su fúnebre crespón;
Pero jamás en mí podrá apagarse
La llama de tu amor

Gustavo Adolfo Bécquer

Quiero dedicar esta tesis doctoral a la persona con quien comparto mi vida...
“A tí, amor mío, por saber enamorarme todos los días, por estar siempre a mi lado, por mimarme y cuidarme, por los sacrificios que has hecho por mí, por tu dedicación, por tu cariño y tu amor, por los millones de detalles que siempre tienes conmigo, por comprenderme, por perdonarme, por los innumerables buenos momentos que me regalas... ”.

A Silvia

Detrás de cada gran hombre hay una gran mujer. Detrás de ella, está su esposa.

- Groucho Marx

Menos mal que no soy un gran hombre y mi mujer no es mi esposa

- Kike

Índice

| | Página |
|---|--------|
| 1. Resumen – Summary | 1 |
| 2. Introducción | 9 |
| 2.1. Clase magistral | 10 |
| 2.2. Aprendizaje basado en problemas | 11 |
| 2.2.1. Aprendizaje basado en la discusión de casos/problema | 14 |
| 2.3. La Mejor Evidencia en Educación Médica | 17 |
| 2.4. Educación basada en resultados | 19 |
| 2.5. Educación basada en competencias | 20 |
| 2.5.1. Evaluación de las competencias | 22 |
| 2.6. Justificación del estudio | 26 |
| 3. Hipótesis y Objetivos | 30 |
| 4. Material y métodos | 32 |
| 4.1. Metodología general | 33 |
| 4.1.1. Diseño | 33 |
| 4.1.2. Sujetos | 33 |
| 4.1.2.1. Cálculo del tamaño muestral | 33 |
| 4.1.2.2. Población del estudio | 34 |
| 4.1.2.3. Criterios de inclusión | 35 |
| 4.1.3. Enseñanza | 36 |
| 4.1.4. Evaluaciones y estudios piloto | 36 |
| 4.1.5. Análisis estadístico | 37 |
| 4.2. Metodología artículo 1. | 38 |
| 4.3. Metodología artículo 2. | 44 |
| 4.4. Metodología artículo 3. | 49 |
| 5. Resultados | 54 |
| 5.1. Resultados artículo 1. | 55 |
| 5.2. Resultados artículo 2. | 59 |
| 5.3. Resultados artículo 3. | 72 |
| 6. Discusión | 79 |
| 7. Conclusiones | 97 |
| 8. Bibliografía | 100 |
| 9. Anexos | 111 |
| 9.1. Artículo 1. Carrero E , Gomar C, Penzo W, Fabregas N, Valero R, Sanchez-Etayo G. Teaching basic life support algorithms by either multimedia presentations or case based discussion equally improves the level of cognitive skills of undergraduate medical students. Med Teach 2009; Feb 24:e1-e7. Disponible en: URL; http://dx.doi.org/10.1080/01421590802512896 | 112 |

| | |
|--|-----|
| 9.2. Artículo 2. Carrero E , Gomar C, Penzo W, Rull M. Comparison between lecture-based approach and case/problem-based learning discussion for teaching pre-anaesthetic assessment. Eur J Anaesthesiol 2007;24:1008-15. | 120 |
| 9.3. Artículo 3. Carrero E , Gomar C, Fábregas N, Penzo W, Castillo J, Villalonga A. Clase magistral versus aprendizaje basado en caso/problema para la enseñanza del embolismo aéreo en formación médica continuada. Rev Esp Anesthesiol Reanim 2008; 55: 202-9. | 129 |
| 10. Comunicaciones a reuniones científicas derivadas de esta tesis | 138 |

1. Resumen

ANTECEDENTES: Existe la necesidad de innovar en educación médica y medir los resultados de la enseñanza. La enseñanza basada en la *clase magistral* sigue siendo el método de enseñanza tradicional más extendido. El método *aprendizaje basado en la discusión de casos/problema* es una variante del método aprendizaje basado en problemas que puede superar sus limitaciones y se adapta bien a los programas de formación *de pregrado, postgrado y Formación Médica Continuada*. Este método es fácil de implementar sin la necesidad de grandes recursos educativos y ha demostrado una gran aceptación entre los participantes. No hay evidencias de cuál es el mejor método de enseñanza de la Anestesiología. La metodología de la enseñanza de la Anestesiología es un campo de investigación poco explorado. Se reconoce que el diseño de métodos de evaluación capaces de medir y comparar la eficacia de la enseñanza es dificultoso. Los métodos de evaluación han de definir con claridad cuales son las competencias que están midiendo. La investigación docente en Anestesiología ha de basarse en la mejor evidencia existente. No hemos encontrado estudios controlados, prospectivos y randomizados que comparen la *clase magistral* y el *aprendizaje basado en la discusión de casos/problema* para la enseñanza de la Anestesiología.

OBJETIVO: Comparar la eficacia del *aprendizaje basado en la discusión de casos/problema* con el método de enseñanza tradicional basado en la *clase magistral* en la enseñanza de tres temas de Anestesiología enmarcados, respectivamente, en tres programas diferentes de formación: *soporte vital básico (pregrado), valoración pre-anestésica (post-grado) y embolismo aéreo intraoperatorio (Formación Médica Continuada)*.

MATERIAL Y MÉTODOS: Hemos realizado tres estudios controlados, prospectivos y randomizados. La población de los estudios incluyó, respectivamente, 68 estudiantes de medicina, 54 residentes de Anestesiología y 52 anestesiólogos expertos. Los alumnos quedaron randomizados para recibir una *clase magistral (grupo clase magistral)* o una sesión de *discusión de casos/problema (grupo caso/problema)* del mismo tiempo de duración. Los objetivos docentes se definieron con antelación y eran los mismos para ambos grupos. Los profesores eran expertos en el tema y en la metodología docente. Los estudios incluyeron una evaluación antes de la enseñanza (*pre*) y otra inmediatamente después (*post*). Las evaluaciones se basaron en situaciones clínicas hipotéticas que cumplían los objetivos docentes. Se realizó una prueba piloto previa a cada estudio. Las variables de las evaluaciones fueron diferentes campos del conocimiento que medían habilidades cognitivas e integrativas. Cada campo incluía un número determinado de *ítems* sobre los que se contabilizó el número total de respuestas correctas. Las evaluaciones fueron puntuadas por dos evaluadores diferentes a los dos profesores, ciegos a la asignación de grupos. Consideramos relevante que el porcentaje de alumnos que mejoraron sus puntuaciones en cada campo fuese superior al 50%. Definimos la significación estadística como $P < 0,05$.

RESULTADOS: Ambos métodos mejoraron de forma similar los campos del conocimiento analizados, aunque esta mejoría fue mínima en el programa de *Formación Médica Continuada*. El índice *kappa* de concordancia entre evaluadores fue > 0.8 para todas las evaluaciones. En el programa de *pregrado* no encontramos diferencias significativas entre los grupos en la *toma de decisiones basada en la secuencia de soporte vital básico* ni en la *capacidad de detectar errores en la*

secuencia de actuación. Con ambos métodos las puntuaciones de las evaluaciones *post* mejoraron por igual y de forma significativa (valores de $P < 0,001$). La mayoría de los estudiantes no obtuvieron la máxima puntuación para ninguna de las evaluaciones *post*. El análisis de los incrementos absolutos de puntuación y la ganancia de puntuación no mostró ninguna diferencia significativa entre los grupos. Tampoco encontramos diferencias entre los grupos en el número de estudiantes que mejoraron sus puntuaciones *post*. En ambos grupos, los porcentajes de estudiantes que mejoraron sus puntuaciones fueron relevantes; estas fueron mayores en la detección de errores (*grupo clase magistral*: 85,3%, *grupo caso/problema*: 82,3%) que en la toma las decisiones correctas (*grupo clase magistral*: 55,9%, *grupo caso/problema*: 58,8%). En el *grupo clase magistral* esta diferencia fue significativa ($P = 0,012$). En el programa de *postgrado*, antes de la enseñanza, la puntuación en el campo 1 “reconocimiento de datos clínicos con implicaciones anestésicas” fue significativamente mayor en el *grupo clase magistral* ($P = 0,006$). Ambos métodos de enseñanza mejoraron las puntuaciones de este campo y de los campos 2 “razonamiento de datos clínicos con implicaciones anestésicas” y 4 “clase Mallampati”; el *grupo caso/problema* también mejoró los campos 3 “clase ASA” y 6 “razonamiento de la elección de la técnica anestésica”; ningún método de enseñanza mejoró el campo 5 “elección de la técnica anestésica”. Después de la enseñanza, la puntuación del campo 1 persistió significativamente mayor en el *grupo clase magistral* ($P = 0,005$) y la puntuación del campo 3 fue significativamente mayor en el *grupo caso/problema* ($P = 0,044$). El análisis de las diferencias entre las puntuaciones *post* y *pre* no reveló diferencias significativas entre los dos grupos para ninguno de los campos medidos. No encontramos diferencias entre los grupos en el número de residentes que mejoraron sus puntuaciones *post*. Los porcentajes de residentes que mejoraron sus puntuaciones en los campos 1, 2 y 4 fueron relevantes en ambos grupos. Encontramos diferencias significativas en ambos grupos al comparar el porcentaje de residentes que mejoró su puntuación entre los distintos campos de conocimiento medidos (valores de $P = 0,000$). Después de la enseñanza, encontramos correlación lineal significativa entre los campos 1 y 2; la correlación fue leve en el *grupo clase magistral* ($r^2 = 0,168$) y moderada en el *grupo caso/problema* ($r^2 = 0,467$). En el programa de *Formación Médica Continuada*, la evaluación *pre* mostró que el conocimiento previo de los participantes sobre el tema ya era elevado. No hubo diferencias significativas entre los dos métodos de enseñanza para ninguno de los campos de conocimiento antes o después de la enseñanza. Los participantes que atendieron la *clase magistral* mejoraron sus puntuaciones en “monitorización” ($P = 0,03$) y “tratamiento” ($P = 0,001$). El *grupo caso/problema* también mejoró sus puntuaciones en “tratamiento” ($P = 0,003$). No se encontraron diferencias significativas entre los dos grupos al comparar la diferencia entre las puntuaciones *post* y *pre* en ninguno de los campos. No encontramos diferencias entre los grupos en el número de participantes que mejoraron sus puntuaciones después de la enseñanza. El porcentaje de participante que mejoró sus puntuaciones en el campo “tratamiento” fue relevante en ambos grupos. Encontramos diferencias significativas en el *grupo clase magistral* al comparar el porcentaje de participantes que mejoró su puntuación entre los distintos campos de conocimiento medidos ($P = 0,005$). Encontramos correlación lineal leve entre el reconocimiento y el razonamiento de los ítems del campo “factores de riesgo y síntomas” en el *grupo caso/problema* antes de la enseñanza ($r^2 = 0,261$).

CONCLUSIONES: La eficacia de la *clase magistral* y el *aprendizaje basado en la discusión casos/problema* fue similar en términos de mejoría inmediata de la adquisición de conocimientos, la capacidad de aplicar el conocimiento para la solución de problemas, el manejo de la información, el razonamiento clínico o unir el conocimiento básico y clínico. El diseño y metodología aplicados permitió una valoración objetiva de las habilidades cognitivas e integrativas para los temas, poblaciones de alumnos y contextos estudiados.

Palabras clave: EDUCACIÓN, educación basada en la evidencia; DISEÑO DE INVESTIGACIÓN; ESTUDIOS PROSPECTIVOS; ENSAYOS RANDOMIZADOS CONTROLADOS; MÉTODOS; MEDIDAS EDUCATIVAS, conocimiento; COMPETENCIAS PROFESIONALES; ANESTESIA; EDUCACIÓN MÉDICA PREGRADO; RESIDENCIA; FORMACIÓN MÉDICA CONTINUADA; ENSEÑANZA, clases magistrales, aprendizaje basado en problemas; REANIMACIÓN CARDIOPULMONAR, soporte vital básico; CUIDADOS PREOPERATORIOS; EMBOLISMO AÉREO.

Summary

BACKGROUND: There is a need for innovation in medical education and measure the outcomes of teaching. *Lecture* remains the most common method of teaching. *Case/problem based learning* is a modified form of the problem based learning method; it may overcome its limitations and it adapts well into the training programs for *undergraduate medical education, internship and residency and continuing medical education*. This method is easy to implement, it does not require great institutional resources and a high degree of satisfaction among students has been addressed. There is currently no evidence on the best anaesthetic teaching method. The methodological aspect of teaching anaesthesiology is an under-explored research field. It is known that the design of assessment methods able to measure and compare teaching effectiveness is difficult. Assessment methods have to exactly define the competences they measure. Educational research on anaesthesiology must be developed based on the best available evidence. As far as we know there are not prospective and randomized controlled studies comparing *lecture* and *case/problem based learning* for teaching anaesthesiology.

OBJETIVE. To compare the effectiveness of the *case/problem based learning* method with that of the *lecture* based approach for teaching three different anaesthetic topics framed in three different training programs: *basic life support (undergraduate medical education), pre-anesthetic assessment (resident training program)* and *air embolism (continuing medical education)*.

MATERIAL AND METHODS. We carried out three prospective and randomized controlled studies. The studies included, respectively, 68 undergraduate medical students, 54 residents of anaesthesiology and 52 expert anaesthesiologists. Participants were randomized to receive a *lecture (lecture group)* or a *case/problem based learning session (case/problem group)*, both lasting the same time. The teaching objectives for both groups were the same and had been previously defined. The participating teachers had teaching experience in the topics and also with the teaching methods applied. Assessments were completed before (*pre*) and immediately after (*post*) teaching; they were based on hypothetical clinical situations fulfilling the teaching contents. A pilot trial was performed earlier for the design of each study. The primary endpoints assessed were different knowledge fields related to cognitive and integrative skills. Each field had a fixed number of *items* for which the total number of correct answers was counted. Assessments were scored by two evaluators, blinded to group allocation, different from the teachers. The percentage of participants who improved their scores in each field were counted; we considered a percentage of over 50% to be relevant. Statistical significance was defined as $P < 0,05$.

RESULTS. Both teaching methods obtained similar improvements for the fields of knowledge measured, although improvements were minimal for the *continuing medical education* program. The *kappa* index inter-evaluator agreement was higher than 0,8 for all assessments. At the *undergraduate medical education* program, we found no significant differences between the groups in “*decisions based on basic life support sequence*” nor in “*detection of errors on basic life support performance*”. In both groups, *post*-scores improved significantly (P values $< 0,001$) at the same grade. Most students did not reach the maximum score at any of the *post*-teaching assessments. The analysis of the actual increases of scores and the gains of scores did not reveal any significant differences between the

groups. We did not find differences between the groups in the number of students who improved their *post*-scores for either assessment. In both groups, the percentages of students who improved their scores were relevant; these percentages were higher on *detection of errors (lecture group: 85,3%, case/problem group: 82,3%)* than on *correct decisions based on basic life support sequence (lecture group: 55,9%, case/problem group: 58,8%)*. In *lecture group* this difference was significant ($P = 0,012$). At the *resident training program*, before the teaching session, the *lecture group* scored significantly higher on field 1 “*recognizing clinical data with anaesthetic implications*” ($P = 0,006$). Both teaching methods improved scores on this field and on field 2 “*reasoning clinical data with anaesthetic implications*” and field 4 “*Mallampati class*”; the *case/problem group* also improved on field 3 “*ASA class*” and field 6 “*reasoning choice of anaesthetic technique*”; no group improved its score on field 5 “*choice of anaesthetic technique*”. After the teaching session, the field 1 score was still significantly higher in the *lecture group* ($P = 0,005$), and the field 3 score was significantly higher in the *case/problem group* ($P = 0,044$). Analysis of the difference between *pre*- and *post*-scores revealed no significant differences between the two groups for any of the fields studied. We found no differences between the groups in the number of residents who improved their *post*-scores. The percentages of residents who improved their scores on fields 1, 2 and 4 were relevant in both groups. We found significant differences, in both groups, when comparing the percentage of residents who improved their scores among the different fields of knowledge measured (P values = 0,000). After the teaching session, we found significant correlation between fields 1 and 2; correlation was weak in the *lecture group* ($r^2 = 0,168$) and modest in the *case/problem group* ($r^2 = 0,467$). At the *continuing medical education program*, *pre*-assessment scores showed a high level of knowledge among participants. There were no significant differences between the two learning groups for any of the knowledge fields before or after teaching. Participants who attended to the *lecture* improved their scores for knowledge of “*monitoring*” ($P = 0,03$) and “*treatment*” ($P = 0,001$). Participants in the *case/problem group* also improved their scores for knowledge of “*treatment*” ($P = 0,003$). The analysis of the actual increases of scores on the fields did not reveal any significant differences between the groups. We found no differences between the groups in the number of participants who improved their scores after teaching. The percentage of participants who improved their scores in the field “*treatment*” were relevant in both groups. There were significant differences in *lecture group* when comparing the percentage of participants who improved their scores among the different fields of knowledge measured ($P = 0,005$). We found a weak correlation between recognition and reasoning of the *items* from field 1 “*risk factors and symptoms*” in the *case/problem group* before teaching ($r^2 = 0,261$).

CONCLUSIONS. The effectiveness of *lecture* and *case/problem based learning* was similar in terms of improving participants’ immediate acquisition of knowledge, capacity for applying knowledge to solve problems, information management, clinical reasoning or linking basic and clinical knowledge. Our design and methodology allowed an objective assessment of these cognitive and integrative skills for the topics, populations and the contexts studied.

Key words: EDUCATION, evidence based education; RESEARCH DESIGN; PROSPECTIVE STUDIES; RANDOMIZED CONTROLLED TRIALS; METHODS; EDUCATIONAL MEASUREMENT, knowledge; PROFESSIONAL COMPETENCES; ANAESTHESIA; UNDERGRADUATE MEDICAL

Summary

EDUCATION; INTERNSHIP AND RESIDENCY; CONTINUING MEDICAL EDUCATION; TEACHING, lectures classes, problem based learning; CARDIOPULMONARY RESUSCITATION, basic life support; PREOPERATIVE CARE; AIR EMBOLISM.

2. INTRODUCCIÓN

Un componente fundamental de los programas docentes en Anestesiología es seleccionar los métodos de enseñanza. El objetivo es que los alumnos procesen mejor la información integrando los conocimientos nuevos a los previamente adquiridos. En educación médica, persiste la controversia sobre qué método de enseñanza es más efectivo, más eficiente en relación a los recursos que precisa y más adecuado al perfil del profesor y del alumno.

2.1. Clase magistral

La enseñanza basada en la *clase magistral* o *lección magistral* ha sido el método tradicional de transmitir los conocimientos teóricos. Es uno de los métodos más antiguos y sin duda el más frecuentemente utilizado [1-3]. El profesor desempeña el papel de transmisor de unos conocimientos que el alumno ha de asimilar; el alumno juega, por tanto, un papel pasivo en el aprendizaje [4]. Su principal ventaja es que garantiza una amplia cobertura del temario a grupos grandes de alumnos sin precisar grandes recursos educativos [3,5]. La *clase magistral* permite al profesor cumplimentar los objetivos docentes de una forma metodológica y organizada y asegura que cada participante recibe la información. Otras ventajas potenciales son: permite introducir temas difíciles, mostrar diferentes perspectivas de un tema, hacer revisiones y puestas al día de las investigaciones, introducir guías clínicas e incorporar experiencias relevantes personales, clínicas o de pruebas complementarias; con un estilo y formato adecuados y con las habilidades necesarias pueden estimular el pensamiento clínico y científico y favorecer la comprensión [3]. La *clase magistral* es, sin duda, el método más eficiente en conseguir el objetivo de distribuir la información; sin embargo, su eficacia ha sido cuestionada [6,7].

Una de las principales críticas de los modelos tradicionales de enseñanza es que los alumnos tienen dificultades para integrar e interpretar los conocimientos nuevos que reciben en las *clases magistrales*, no piensan de manera crítica, retienen poca información y fracasan a la hora de incorporar estos conocimientos para resolver problemas de la práctica clínica [8,9]. Que los alumnos reciban los objetivos docentes no garantiza que los hayan aprendido. Por este motivo, se han recomendado cambios profundos en la educación médica que conduzcan a unas estructuras conceptuales mejores y promuevan habilidades que favorezcan un aprendizaje duradero.

La educación está cambiando el enfoque de la enseñanza al aprendizaje (Tabla 1). Los cambios principales promulgan un aprendizaje activo y basado en problemas, con participación del alumno en su proceso educativo [8]; un aprendizaje que estimule el pensamiento crítico, es decir, la síntesis, análisis y evaluación de la información, que conduce al alumno a tomar decisiones y que permita comprender mejor la actuación del alumno o profesional en una situación realística. En los últimos años se han introducido nuevas técnicas docentes con el objetivo de mejorar el aprendizaje, las competencias y la excelencia profesional: simuladores, enseñanza on-line, seminarios, grabaciones en video y el aprendizaje basado en problemas. En el momento actual no existe un consenso sobre cuál es el mejor método de enseñanza.

Tabla 1. Transformación de la enseñanza al aprendizaje.

| De la enseñanza | Al aprendizaje |
|---|--|
| Transferencia del conocimiento del profesor al alumno | El alumno descubre y construye el conocimiento |
| Ofrece clases/cursos | Crea un entorno de aprendizaje |
| Calidad de la enseñanza | Calidad del aprendizaje |
| Contenidos | Resultados |
| Acceso a la educación | Éxito del aprendizaje |
| Aprendizaje variable; tiempo constante | Tiempo variable; aprendizaje constante |
| Aprendizaje acumulativo y lineal | Aprendizaje anidado e interactivo |
| Clase competitiva, individualista | Entorno de aprendizaje cooperativo, soporte de grupo |
| El profesor y el alumno actúan independientes | El profesor y el alumno trabajan en equipo |
| Asignaturas independientes | Temario cruzado |
| El conocimiento existe en alguna parte | El conocimiento existe en la mente del alumno |

Tomado de Tomey AM [10].

2.2. Aprendizaje basado en problemas

El *aprendizaje basado en problemas* empezó a utilizarse en la enseñanza de medicina en la Universidad de McMaster (Ontario, Canadá) en 1969 [11,12] y ha sido la mayor innovación en educación médica de los últimos años. Desde entonces lo han incorporado muchas facultades en todo el mundo [13,14] y actualmente sigue siendo objeto de numerosas publicaciones [15-18]. Hay cuatro principios clave de la teoría de la educación moderna que pueden ser aplicados al *aprendizaje basado en problemas*: constructivismo, auto-aprendizaje, colaboración y aprendizaje contextual [19].

El *aprendizaje basado en problemas* se inicia con la presentación a un grupo pequeño de alumnos de un problema, habitualmente una situación clínica como punto de partida de la discusión. Se plantea un problema que reta el conocimiento y experiencia de los alumnos. A lo largo de varias sesiones, los alumnos identifican sus propios objetivos de aprendizaje, disponen de tiempo para explorarlos; salen, buscan individualmente la información y se reúnen de nuevo para compartirla y discutirla con el resto del grupo. A través del esfuerzo autodidacta de los alumnos, el grupo integra los aspectos científicos y clínicos de cada tema. Finaliza con una discusión en grupo y auto evaluación de la actuación de los alumnos [18,20,21]. En el *aprendizaje basado en problemas* el profesor actúa exclusivamente como facilitador, no interviene en las discusiones ni transmite conocimientos; su función es observar el progreso del grupo y guiarlo a través de su propio aprendizaje [22]. El profesor se asigna para promover el desarrollo del razonamiento científico, el auto estudio y la auto evaluación [23].

El *aprendizaje basado en problemas* considera el aprendizaje como un proceso de construcción activa del conocimiento más que como una transmisión o recepción pasiva del mismo; la activación de los conocimientos previos es esencial, permitirá al alumno integrar la información nueva sobre el conocimiento existente; la resolución del problema no es el fin perseguido en sí mismo [18], el *aprendizaje basado en problemas* enfatiza el proceso de resolución del problema como medio que

permitirá a los alumnos obtener la información relevante que les conduzca a alcanzar los objetivos docentes [22]. A diferencia de la *clase magistral* enfatiza la creación del conocimiento más que la adquisición del mismo. El *aprendizaje basado en problemas* estimula el auto aprendizaje y el trabajo en equipo [24]. El aprendizaje autodirigido es un elemento esencial [18]. Los alumnos han de aprender a buscar la información, compartirla, analizarla, utilizarla e integrar el conocimiento. Los alumnos aprenden una metodología de formación, juegan un papel activo en su aprendizaje; el aprendizaje se centra, por tanto, en el alumno. Algunos autores han definido el *aprendizaje basado en problemas* como modelo de aprendizaje abierto [21].

Sus defensores argumentan que favorece la retención del conocimiento, simula la práctica clínica, estimula la curiosidad científica y favorece una amplia comprensión de la complejidad de la medicina [25-28]. Otras ventajas potenciales incluyen: fomentar habilidades de aprendizaje que serán útiles toda la vida profesional, mayor motivación, pensamiento crítico, juicio clínico, trabajo en equipo, habilidades comunicativas, relaciones interpersonales, integración del conocimiento en la práctica, auto-evaluación y desarrollo de un aprendizaje más constructivo [22,29].

Esta filosofía del aprendizaje cambia radicalmente la presunción tradicional de que la información ha de adquirirse antes de que se plantee la necesidad de solucionar un problema y que el aprendizaje ha de ser secuencial progresando desde los conceptos básicos hasta su aplicación clínica. El *aprendizaje basado en problemas* ofrece un entorno en el que los alumnos pueden progresar más allá de sus conocimientos previos, aprender dentro de un contexto real y reforzar sus conocimientos gracias al trabajo en grupos pequeños. El *aprendizaje basado en problemas* es una forma de aprendizaje contextual [30]. Aunque no formaron parte del diseño original, a Neufeld y Barrows [23] se le reconoce el haber definido y popularizado el método clásico de *aprendizaje basado en problemas* en educación médica.

El *aprendizaje basado en problemas* tiene limitaciones [29,31,32]. El conocimiento se adquiere de una forma menos organizada [29]. Se le ha considerado un método incapaz de transmitir una visión comprensiva de las ciencias básicas [32], incapaz de cumplimentar, por tanto, todo el programa docente. Su implementación está condicionada por los importantes recursos humanos y económicos que precisa: en primer lugar, el tiempo de dedicación de los profesores y alumnos es mayor [33]. Los profesores han de disponer de tiempo libre para atender las consultas de los alumnos y tiempo para la planificación, organización y puesta en marcha del método. Los resultados más optimistas estiman que el *aprendizaje basado en problemas* incrementa un 30% el tiempo de programación del profesorado [34,35] frente a los que afirman que el tiempo de dedicación del profesorado es 3-4 veces mayor que la *clase magistral* [36]. Los alumnos precisan más tiempo para desarrollar el auto aprendizaje y este consumo de tiempo excesivo puede producir frustración entre los alumnos y profesores [37,38]. En segundo lugar, el *aprendizaje basado en problemas* resulta más caro; se necesitan más salas para las actividades de los grupos y los costes económicos se disparan si el número total de alumnos es de más de 100 [36].

Un fenómeno bien reconocido asociado al *aprendizaje basado en problemas* es la disfunción del grupo [37,39,40]. La falta de medios alternativos de aprendizaje (*clases magistrales*, apuntes editados

etc) puede estresar al alumno, hacerlo escapar del grupo e intentar sobrevivir aprendiendo de memoria. Una disfunción importante es el denominado comportamiento ritual, que implica indiferencia hacia la discusión del grupo y/o fracaso en la preparación o asistencia regular a las sesiones.

Shanley [32] afirma que los grupos en el *aprendizaje basado en problemas* no funcionan por tres razones: las diferencias inevitables en las habilidades, motivaciones y experiencia entre los profesores; las discusiones poco enriquecedoras por el conocimiento limitado de los alumnos en el tema y el conflicto en el grupo entre la ética del individualismo y el espíritu del colectivismo. Este autor considera el conocimiento y experiencia previa requisitos esenciales para que el *aprendizaje basado en problemas* sea una experiencia enriquecedora. Los alumnos tienen que estar familiarizados con el tema porque así identificarán mejor cuáles son sus carencias y cómo resolverlas. El auto-aprendizaje no ha de interpretarse como auto-enseñanza. Cobb [41] lamenta que “*el auto-aprendizaje se transforme frecuentemente en que los profesores no deben decirles nada a los alumnos porque ellos deben construir su propio conocimiento*”. Schmidt [42] sugiere que los alumnos necesitan un mínimo nivel de estructura clínica para aprender y, si no la tienen, precisan la guía del profesor. Existe poca garantía de que los alumnos aprendan como aplicar el conocimiento necesario para la práctica clínica en ausencia de una dirección clínica apropiada [43].

Muchos profesores se sienten infravalorados y desmotivados actuando sólo como facilitadores [32]. Existe un sentimiento de aprehensión entre el profesorado, incluso entre los más defensores del *aprendizaje basado en problemas*, porque no existe la oportunidad de incluir conocimientos basados en experiencias clínicas cuando se introduce esta metodología en asignaturas clínicas [38]. La importancia y necesidad de que el facilitador sea también un profesor experto en el tema como determinante del buen funcionamiento del grupo ha sido referenciada en la literatura [44,45]. Por otro lado, es difícil formar profesores para la habilidad de ser facilitadores más que para las habilidades tradicionales de enseñar [29,32]. Otro de los riesgos potenciales del *aprendizaje basado en problemas* es que se acaben construyendo casos demasiado artificiales que despisten al alumno [32].

Por último, comparado con el formato tradicional de enseñanza el *aprendizaje basado en problemas* ha demostrado mejorar la satisfacción de los alumnos y su capacidad de enfrentarse a los problemas [46,47] pero no mejorar los conocimientos o las habilidades técnicas [16,48]. A pesar de lo atractivo de su filosofía, la superioridad del *aprendizaje basado en problemas* no ha podido ser demostrada de manera fehaciente [49,50]. Estas limitaciones crean incertidumbre sobre el futuro de esta metodología docente.

2.2.1 Aprendizaje basado en la discusión de casos/problema

Diversos formatos del aprendizaje basado en problemas han surgido [51], en la mayoría de casos, como respuesta a la necesidad de adaptar este método docente a los programas de enseñanza y que se alejan progresivamente del modelo curricular de aprendizaje basado en problemas [28,52]. El *aprendizaje basado en la discusión de casos/problema* [25,53-55] también llamado *aprendizaje basado en casos o en casos/problema* [21,45,56,57] o *método de discusión de casos o casos/problema* [10,58-62] es un método que puede superar las limitaciones del aprendizaje basado

en problemas manteniendo aún un formato interactivo, encaja con los modelos constructivos de aprendizaje y favorece un aprendizaje duradero a lo largo de la *Formación Médica Continuada*. En opinión de Shanley [32] “...liberado de las constricciones de la metodología del aprendizaje basado en problemas, conlleva un nivel de disertación más profunda con un coste mucho menor”. Algunas publicaciones recientes muestran un interés renovado en este método de enseñanza [21,58-60,62].

El *aprendizaje basado en la discusión de casos/problema* nace de la discusión de casos legales en la Facultad de Derecho de Harvard y se utilizó por primera vez como método docente en 1930 [63]. En educación médica también se ha empleado [58-60,64,65] aunque no de forma tan generalizada. Sin embargo, es un hecho conocido que la mayoría de los médicos tienen la experiencia de aprender de la discusión de casos, como por ejemplo los presentados en sesiones clínicas o publicados en revistas relevantes [66].

El *aprendizaje basado en la discusión de casos/problema* puede considerarse una modalidad del aprendizaje basado en problemas porque comparte características docentes comunes [67], sin embargo, este método tiene aspectos diferenciales que lo convierten en un método de enseñanza propio. Este método cambia del aprendizaje autodirigido por el propio alumno al de un aprendizaje guiado por el profesor pero centrado en el alumno [56]. Proporciona un proceso participativo que incentiva la curiosidad del alumno, favorece un aprendizaje activo y reflexivo y mejora la capacidad de aplicar en la práctica clínica los conocimientos adquiridos gracias a un enfoque más sistemático del problema. [10]. Una de las diferencias fundamentales es que este método requiere que los alumnos tengan algún grado de conocimiento previo que utilizaran para discutir y resolver los casos clínicos [68].

El *aprendizaje basado en la discusión de casos/problema* se centra en el problema que tienen que ser resuelto. Existen una serie de principios para que su diseño sea más efectivo [29,69]. Los casos han de mostrar situaciones clínicas reales. Su presentación debe incluir todos los detalles importantes; el alumno ha de identificar los puntos clave de cada caso y tomar decisiones en función de los dilemas que se plantean. Los alumnos discutirán el caso individualmente o en pequeños grupos y para ello pueden utilizar el conocimiento que previamente han adquirido de *clases magistrales* y otras fuentes de información; de esta manera, el alumno tiene la oportunidad de familiarizarse y profundizar sus conocimientos sobre los temas que surjan en la discusión [10]. El grupo no puede ser demasiado pequeño, tiene que tener un tamaño mínimo para favorecer la discusión. La distribución de los alumnos en semicírculo en la clase también puede estimular su participación [10].

Las discusiones son una de las estrategias más efectivas para promover el aprendizaje activo y el análisis crítico [70]. La discusión está dirigida por el profesor, que además de ser experto en la materia, ha de tener habilidades docentes en el manejo del grupo: saber dirigir la discusión de forma efectiva, saber escuchar, estimular la participación de los alumnos y el intercambio de opiniones, detectar señales que favorezcan la discusión, animar y ayudar a los alumnos a detectar y corregir dudas, conceptos erróneos o déficits de conocimiento, establecer refuerzos positivos, críticas constructivas, control del tiempo etc [58]. La retroalimentación es un instrumento de trabajo esencial,

favorece la comprensión y la corrección de errores por parte del alumno y es un mecanismo esencial para que el profesor imparta los objetivos docentes [71,72]. La retroalimentación que ejerce el profesor es un factor conocido que mejora el aprendizaje [73].

El profesor puede redirigir y explorar los fundamentos incorrectos, dar información correctiva y discutir opciones alternativas [21,25,51]. En opinión de algunos expertos [9], este papel central y activo del profesor en respuesta a las necesidades de los alumnos constituye la pieza clave para el éxito de este método [32]; con este método el profesor ayuda a los alumnos a descubrir y construir el conocimiento en vez de solamente transferirlo [10]. El método *aprendizaje basado en casos/problema* puede adoptar, a su vez, múltiples modificaciones en función del mayor o menor grado de protagonismo del profesor y del alumno [57,58].

La construcción del caso y su calidad narrativa son aspectos fundamentales que condicionan el éxito de este método [29,69,74,75]. Existen una serie de principios para que el diseño de los casos sea más efectivo [29,69]. El caso tiene que ser lo más real posible, despertar curiosidad para estimular la discusión y al mismo tiempo, cubrir los objetivos docentes programados. Es recomendable que el profesor haya vivido el caso directamente porque así se dispondrá de toda la información, detalles adicionales y curso final. El caso no tiene que ser especialmente largo, complicado o raro pero sí debe estar abierto a la posibilidad de soluciones alternativas [62]. El formato del caso puede variar (escrito, video, electrónico, "role-plays") [10,56,76] y puede entregarse con antelación [58,60] o en la misma clase [45,77]; el caso también suele incluir preguntas abiertas para facilitar la discusión [10].

El profesor ha de estar formado y entrenado en el uso de este método de enseñanza [60]. Los casos pueden ser re-utilizados más veces siempre que sean todavía apropiados. De hecho se recomienda su re-utilización porque permite que el profesor se beneficie de la experiencia adquirida, incluyendo la evaluación de los alumnos, con el consiguiente ahorro de tiempo de preparación [62]. Es importante disponer de un tiempo final de reflexión, intercambio de opiniones y resumen que permita al alumno extraer conclusiones y futuras aplicaciones prácticas de lo aprendido [10].

El *método aprendizaje basado en la discusión de casos/problema* precisa menos recursos docentes que el aprendizaje basado en problemas. El tamaño del grupo puede ser varias veces mayor, la enseñanza se completa en una sola sesión, mantiene un tiempo de programación del profesorado en unos niveles equiparables a los formatos de enseñanza tradicional basados en la *clase magistral*, precisa menos tiempo de dedicación del alumno [21]. Debido a su enfoque del aprendizaje más dependiente del profesor, este método es también menos susceptible de la disfunción del grupo [62]. El *aprendizaje basado en la discusión de casos/problema* es un método sencillo de implementar, con un alto grado de aceptación entre los alumnos y que no precisa costes ni consumo de tiempo adicionales; de hecho desde 1991 la *American Society of Anesthesiologists (ASA)* viene incluyendo el aprendizaje basado en la discusión de casos/problema en sus cursos de *Formación Médica Continuada* [53]. Es un método fácil de poder incorporar como complemento a cualquier sistema curricular [32].

Un estudio reciente que compara las percepciones de los alumnos y profesores entre el aprendizaje basado en problemas y el *aprendizaje basado en la discusión de casos/problema* en dos facultades

de medicina mostró que más del 80% de los alumnos y profesores en ambas universidades prefirieron el *aprendizaje basado en la discusión de casos/problema* por tres motivos principales: menos posibilidad de irse por las ramas (conocimiento disperso), menos carga de trabajo y más oportunidades para la aplicación clínica de las habilidades adquiridas [21]. Se trata de un método simple y fácil de practicar, con una proporción profesor/alumno más flexible que el aprendizaje basado en problemas lo que favorece la persistencia del método a largo plazo [58]. Este método integra mejor el aprendizaje nuevo con el conocimiento previo del alumno. Su diseño ayuda a que se reflejen mejor los objetivos docentes de un determinado tema [78].

Al igual que el aprendizaje basado en problemas, el *aprendizaje basado en la discusión de casos/problema* favorece la comprensión más que la memorización, la aplicación del conocimiento en la resolución de problemas, el pensamiento crítico y la participación del alumno en su aprendizaje [10,57,79]. Otras ventajas potenciales que se le ha atribuido es que los alumnos aprecian mejor el complejo proceso de toma de decisiones clínicas y del cuidado del paciente, lo que puede serles particularmente útil antes de enfrentarse a la realidad clínica [57,80]. La Tabla 2 muestra las principales diferencias entre el aprendizaje basado en problemas y el *aprendizaje basado en la discusión de casos/problema*.

Los problemas potenciales del *aprendizaje basado en la discusión de casos/problema* incluyen la motivación de los alumnos, las habilidades del profesor y el entorno de la enseñanza. No todos los alumnos participan por igual y algunos tienen el riesgo de convertirse en alumnos pasivos; no todos los profesores tienen las mismas capacidades para manejar la discusión y el grupo; los profesores tienen que estar entrenados en esta metodología; el entorno físico de la enseñanza ha de ser propicio y el material de los casos y su preparación precisa tiempo y recursos [10].

Desde hace unos años nosotros hemos introducido el *aprendizaje basado en la discusión de casos/problema* para determinados contenidos de los programas docentes de Anestesiología, sin embargo, a pesar de la motivación del profesorado y su buena acogida entre los alumnos, desconocíamos si aplicando este nuevo método los alumnos aprendían más que con nuestro método tradicional de enseñanza basado en la *clase magistral*.

Aunque existen fundamentos para creer que estas nuevas estrategias docentes deben conducir a mejorar los resultados del aprendizaje comparado con el método tradicional de la *clase magistral*, son escasos los estudios publicados que comparen la *clase magistral* y el *aprendizaje basado en la discusión de casos/problema* utilizando el mismo contenido y los mismos alumnos. Esta falta de datos acontece en un momento en el que desde las principales instituciones de la educación médica se enfatiza la necesidad de cambiar desde la enseñanza basada en la opinión hacia una enseñanza que se base más en la evidencia. [81].

Tabla 2. Características del aprendizaje basado en problemas (ABP) y el aprendizaje basado en la discusión de casos/problema (ABDCP).

| Rasgo | ABP | ABDCP |
|--------------|--|--|
| Presentación | Problema o caso clínico | Problema o caso clínico |
| Tema | No revelado con anterioridad. No posibilidad de preparación previa. | Revelado con anterioridad. Posibilidad de preparación previa. |
| Principio | Proceso de búsqueda del conocimiento dirigido por el grupo de alumnos. Auto-aprendizaje o aprendizaje abierto. Aprendizaje no estructurado. | Solución del problema dirigido por el profesor. Aprendizaje dirigido o guiado. Aprendizaje estructurado. |
| Estrategia | Definir objetivos docentes (alumnos). Tormenta de ideas: grupo. Buscar información, recopilarla, compartirla, discutirla, integrarla (alumnos). Profundizar el conocimiento (alumnos) Transmitir y explicar información (alumnos). No retroalimentación (profesor). | Objetivos docentes predefinidos (profesor). Tormenta de ideas: alumno. Información pre-existente. Plantear el caso, prepararlo, sugerir soluciones, discutirlo, explorar y resolver dudas, alcanzar acuerdos (alumnos y profesor). Profundizar el conocimiento (alumno) Transmitir y explicar información (profesor). Retroalimentación (profesor). |
| Fuentes | Prospectivas: biblioteca, internet, entrevistas. | Retrospectivas: clases magistrales, material docente, libros. |
| Grupo | 5-7 alumnos (< 8). | 15-30 alumnos (<60). |
| Profesor | Facilitador, tutor. Observador. No guía la discusión. No formula ni responde preguntas. No corrige errores. No interactúa | Director, modulador. Participación activa. Guía la discusión. Formula y responde a preguntas. Corrige errores. Interactúa. |
| Sesiones | Múltiples. Requiere trabajo adicional después de la sesiones. | Única. No requiere trabajo adicional después de la sesión. |
| Limitaciones | Grandes recursos económicos y humanos. Profesorado experto. Consumo de mucho tiempo de alumnos y profesores. Disfunción del grupo. No evidencia de mayor eficacia. | Profesorado experto. Construcción adecuada de casos. Riesgo del profesor de caer en clase magistral. Riesgo de pasividad del alumno. Faltan estudios de eficacia. |

Adaptado de Srinivasan M et al. [21] y Tärnvik A [62].

2.3. La Mejor Evidencia en Educación Médica

Existe la necesidad de mejorar la imagen y el valor de la enseñanza. *La Mejor Evidencia en Educación Médica* (“Best Evidence Medical Education” BEME) [82,83], página web: www.bemecollaboration.org, fue un nuevo enfoque que se inició en 1999 con la esperanza de mejorar la enseñanza y el aprendizaje involucrando al profesorado en el proceso de decisión. *La Mejor Evidencia en Educación Médica* puede representarse como un continuo entre el 100% de educación basada en la opinión en un extremo del espectro, donde no es posible encontrar evidencia útil y el 100% de la educación basada en la evidencia en el otro extremo, donde las decisiones pueden tomarse en base a la existencia de una evidencia detallada (Figura 1).

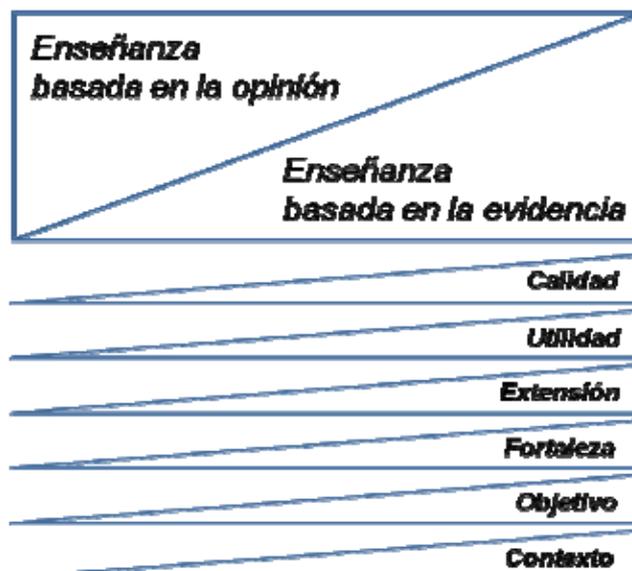


Figura 1. La Mejor Evidencia en Educación Médica.

La búsqueda de la evidencia ha de tener en cuenta una serie de factores o dimensiones [83]: a) *Calidad* o fiabilidad de la evidencia: si bien los estudios de investigación experimental o cuasi-experimentales (estudios randomizados y controlados) constituyen un referente de calidad, en educación médica también es posible obtener fuentes de evidencia basada en la experiencia profesional o el juicio profesional. En investigación docente no hemos de infravalorar la importancia de los métodos de estudio cualitativos [84-86]; b) *Utilidad* o capacidad de aplicar la intervención en otra situación sin adaptaciones; c) *Extensión*: evidencia basada en opiniones, revisiones, meta-análisis etc; d) *Fortaleza* de la evidencia: la evaluación de la evidencia exige ser críticos con los resultados que sólo tienen mínima significación estadística o poca transcendencia docente; e) *Objetivo* o validez de la evidencia: uno de los factores más importantes en la evaluación de la evidencia es conocer el nivel de resultados y de evaluación de la investigación, es decir la calidad del beneficio que buscan las intervenciones docentes; f) *Contexto*: las diferentes fases de la educación, las distintas disciplinas de la misma profesión, la distribución por edad y sexo o diferencias culturales y geográficas pueden influir en los resultados y limitar su generalización. Existe el riesgo de que estas dimensiones no sean coincidentes entre sí. Aunque la situación ideal es que la evidencia disponible se sitúe a la derecha de todas las dimensiones indicadas, ésta puede situarse en cualquier punto entre los extremos de cada uno de estas dimensiones (Figura 1).

El objetivo de la *Mejor Evidencia en Educación Médica* es proporcionar una herramienta de información relevante y accesible que pueda servir de ayuda a los profesores a la hora de tomar decisiones educativas prácticas según la mejor evidencia disponible en cualquier punto de este continuo en el que ellos se encuentren [83]. La *Mejor Evidencia en Educación Médica* es una actitud mental, una filosofía que estimula al docente a que se cuestione su práctica, busque la mejor evidencia disponible, relacione la evidencia con su propia situación y aplique su juicio profesional. Este enfoque estimula el pensamiento crítico ante las prácticas docentes habituales y novedosas en

educación médica y ante la literatura publicada. También ayuda a identificar los déficits y necesidades de investigación futura [87].

Saber si la enseñanza es eficaz no es sencillo de responder. La eficacia de la enseñanza no se mide en términos de sí o no. Una intervención docente puede ser más o menos eficaz dependiendo de qué se enseñe, cómo, cuando, a quién y dónde. Este es el primer filtro de calidad de que han de tener los estudios de investigación. El segundo filtro es el de evaluación. La eficacia tiene que definirse en términos de resultados o competencias adquiridas. Para fundamentar la evidencia son necesarios diseños de estudio prospectivos, randomizados y comparativos, con un tamaño muestral y potencia adecuados, que utilicen un método evaluador que defina con claridad los criterios de eficacia de la enseñanza y que sea a la vez fiable (concordancia entre el mayor número de evaluadores posible), válido (adecuado a la competencia que se quiere evaluar) y reproducible, y que los resultados obtenidos tengan aplicabilidad práctica. El propósito de la evaluación ha de ser claro y su diseño debe estar en concordancia con los objetivos docentes; además, la evaluación debe establecer con antelación cual es el nivel mínimo de competencia que debe alcanzar el alumno [88]. Sin embargo, se reconoce que el diseño de métodos de evaluación que sean capaces de medir y comparar la eficacia de la enseñanza no es fácil [89]. El problema principal al evaluar la enseñanza es la distancia existente entre la educación y su último objetivo, la mejoría de la asistencia sanitaria.

2.4. Educación basada en resultados

Actualmente existe una tendencia hacia la *educación basada en los resultados* [90,91]. Este enfoque se caracteriza por definir con claridad los objetivos docentes que han de alcanzar los alumnos y el diseño de las estrategias docentes y métodos de evaluación para la consecución de dichos objetivos. Los objetivos del aprendizaje han de ser relevantes, claros, factibles y evaluables [18]. De esta manera el alumno tiene claro lo que se espera que aprenda y le ayuda a que su aprendizaje sea más eficiente. La *educación basada en los resultados* también puede ayudar al profesor a planificar mejor el contenido docente, seleccionar la mejor estrategia de enseñanza y evaluación, favorecer la puntuación por objetivos cumplidos, informar mejor de lo que está haciendo, en definitiva lo ayudará a decidir mejor lo que debe enseñar y evaluar. Se ha argumentado que de este modo es probable que mejore la calidad de la educación [92].

Existen distintos modelos para clasificar y analizar los resultados de la enseñanza. El modelo de Kirkpatrick [93] describe 4 niveles: reacción, aprendizaje, comportamiento y resultados y es el que ha adoptado con modificaciones la BEME Collaboration (Tabla 3).

Tabla 3. Modelo de Kirkpatrick para la evaluación de los resultados de la enseñanza.

| | | |
|-----------------|--|--|
| <i>Nivel 1</i> | Reacción | Opiniones de los alumnos sobre la experiencia docente, su organización, presentación, contenido, métodos de enseñanza y calidad |
| <i>Nivel 2A</i> | Aprendizaje. Cambio en actitudes | Cambio en las actitudes o percepciones entre los alumnos hacia la enseñanza y el aprendizaje |
| <i>Nivel 2B</i> | Aprendizaje. Modificación de los conocimientos o habilidades | Para conocimientos, se refiere a la adquisición de conceptos, procedimientos o principios; para habilidades, se refiere a la adquisición de razonamiento, solución de problemas, habilidades psicomotoras y sociales |
| <i>Nivel 3</i> | Comportamiento. Cambio en el comportamiento | Documenta la transferencia del aprendizaje al lugar de trabajo o el deseo de los alumnos de aplicar los nuevos conocimientos y habilidades |
| <i>Nivel 4A</i> | Resultados. Cambio en la práctica a nivel organizativo | Se refiere a los cambios generales en la organización, atribuibles al programa educativo |
| <i>Nivel 4B</i> | Resultados. Cambio en la práctica a nivel de estudiantes, residentes o especialistas | Se refiere a los cambios en la práctica profesional como consecuencia directa de la intervención docente |

Modelo modificado por Freeth et al. [94] y por la BEME Collaboration [95].

2.5. Educación basada en competencias

Los contenidos docentes de la Anestesiología abarcan los tres niveles de formación en educación médica: *pregrado*, *post-grado* y *Formación Médica Continuada*. Los objetivos de la enseñanza son diferentes en cada nivel de acuerdo a las competencias profesionales que han de alcanzar los estudiantes de medicina, los residentes de Anestesiología o los anestesiólogos en cada tema. La *educación basada en las competencias* surgió en los años 50 desde las universidades americanas que demandaban mejores resultados. Cuatro características distinguen la *educación basada en las competencias*: los resultados de los alumnos han de basarse en el análisis de las responsabilidades y funciones de los profesionales; el curriculum debe enfocarse en lo que los alumnos necesitan aprender para ejercer sus responsabilidades profesionales; la estructuración por módulos secuenciados ha de permitir a los estudiantes progresar a su ritmo a través del curriculum y las técnicas de evaluación han de medir la actuación de los alumnos en contextos que se aproximen a la realidad [96].

Existen diferentes factores, además, que demandan que se evalúen las competencias [97]: regularización de la educación médica, calidad de los profesionales e interés social por la seguridad del paciente. Nuevos métodos de evaluación han ido surgiendo simultáneamente a los nuevos enfoques docentes. Las tendencias recientes en formación y evaluación en Anestesiología se han centrado en la adquisición y demostración de las competencias [98,99]. Pero, ¿cuáles son estas competencias para cada nivel de formación? y ¿cuáles son los métodos para evaluarlas? Las

respuestas a estas preguntas no son fáciles de contestar. Es muy difícil saber si un método de enseñanza será más efectivo en formar mejores profesionales en la práctica clínica. Son muchas las influencias posteriores que moldearán la práctica profesional de estos alumnos.

El concepto de competencia es complejo porque debe incluir todos los aspectos transcendentales de la práctica médica. Existen muchas definiciones de competencia [100,101]. El término competencia en su acepción académica se ha definido como la combinación de atributos que describe cómo los estudiantes serán capaces de desenvolverse al finalizar el proceso educativo [102]. Muchos sistemas curriculares definen la competencia profesional en términos de conocimiento, habilidades y actitudes [103], sin embargo, esta definición puede resultar demasiado simplista.

En 1999, *the Accreditation Council for Graduate Medical Education (ACGME)* definió 6 competencias específicas que podrían aplicarse a la formación de residentes: *the ACGME Outcome Project* (www.acgme.org/Outcome). Las competencias eran: atención al paciente (incluía también el razonamiento clínico), conocimientos médicos, aprendizaje basado en la práctica y perfeccionamiento (incluía el manejo de la información), habilidades interpersonales y de comunicación, profesionalismo y práctica basada en sistemas (incluía la economía de la salud y el trabajo en equipo). Epstein y Hundert [104] propusieron posteriormente una definición que ampliaba el concepto de competencia profesional más allá de estas 6 competencias y que engloba la mayor parte de los aspectos de la práctica clínica profesional; definen competencia como *“el uso juicioso y habitual de la comunicación, el conocimiento, las habilidades técnicas, el razonamiento clínico, las emociones, los valores y la reflexión en la práctica diaria para el beneficio del individuo y de la comunidad a la que sirve el profesional”*. La competencia es un hábito que engloba todas las facultades humanas y que se desarrolla a lo largo del tiempo. La competencia ha de entenderse como un todo (competencia general) que engloba sus distintos componentes (competencias específicas, dimensiones o habilidades) [96,104]. Según Epstein y Hundert [104], la competencia profesional incorpora 7 dimensiones o habilidades profesionales: cognitivas; técnicas; integrativas; contextuales; de relación; afectiva/moral y de pensamiento (Tabla 4).

En una revisión reciente, Koh GC et al. [16] adoptan con modificaciones esta clasificación: mantienen la dimensión “técnica”, combinan las dimensiones “de relación” y “cognitivo/moral” por un lado y las “contextuales” y “de pensamiento” por otro, creando dos dimensiones nuevas: “social” y “directiva”, incorporan la dimensión “integrativa” dentro de la “cognitiva” y consideran el “conocimiento”, la “enseñanza” y la “investigación” como competencias independientes; finalmente añaden una última competencia “global”. En nuestra opinión, las 8 dimensiones resultantes no resultan tan comprensibles como las de Epstein y Hundert ni se equiparan tan bien a los niveles de evaluación de las competencias que describimos en el próximo apartado.

En España se ha elaborado recientemente un documento que recoge cuales son las competencias que han de tener los estudiantes de medicina en el campo de la Anestesiología [105]. Una de las competencias implica que los alumnos de medicina han de saber aplicar y enseñar los conocimientos adquiridos sobre *soporte vital básico* y desarrollar las habilidades para su correcta ejecución. Respecto a las competencias profesionales de la formación de los residentes de Anestesiología y de

la *formación continuada* en anestesia se han creado grupos de trabajo entre el Ministerio de Educación y varias sociedades de Anestesiología y Reanimación del estado español y actualmente están en fase de elaboración. En el presente estudio nosotros hemos aplicado la terminología de acuerdo a la definición de Epstein y Hundert [104] por ser la más referenciada en la literatura y la que mejor se adapta a nuestro diseño de estudio.

Tabla 4. Habilidades o dimensiones de las competencia profesional.

| | |
|-------------------------------------|--|
| Habilidades cognitivas | Conocimientos teóricos y de comunicación básica. Manejo de la información. Aplicación del conocimiento a situaciones reales. Uso del conocimiento tácito y las experiencias personales. Solución de problemas abstractos. Adquisición auto-dirigida de nuevos conocimientos. Reconocer déficits de conocimientos. Generar preguntas. Utilización de recursos de aprendizaje. Aprender de la experiencia. |
| Habilidades técnicas | Habilidades en la exploración física. Habilidades en procedimientos. Habilidades quirúrgicas. |
| Habilidades integrativas | Aplicar un juicio científico, clínico y humanístico. Utilizar el razonamiento clínico de forma adecuada. Unir los conocimientos básicos con los clínicos. Manejar las dudas. |
| Habilidades contextuales | Escenario clínico. Uso del tiempo. |
| Habilidades de relación | Habilidades de comunicación. Manejo de conflictos. Trabajo en equipo. Enseñar a otros. |
| Habilidades afectivo/morales | Tolerancia a la ambigüedad y la ansiedad. Inteligencia emocional. Respeto al paciente. Sensibilidad frente al paciente y a la sociedad. Humanismo. |
| Hábitos de pensamiento | Análisis de los propios pensamientos, emociones y técnicas. Atención. Curiosidad crítica. Reconocimiento y respuesta a los sesgos cognitivos y emocionales. Actitud de reconocimiento y corrección de los errores. |

Versión adaptada de la clasificación de Epstein y Hundert [104].

2.5.1. Evaluación de las competencias

El modelo más ampliamente utilizado para la evaluación de las competencias es el de la *pirámide de Miller* [106]. Miller concibe la competencia como una pirámide de 4 niveles (Figura 2). La base de la pirámide consiste en el conocimiento teórico “*saber*”. Un nivel más arriba Miller describe la habilidad de usar el conocimiento en un determinado contexto como “*saber como*”. Este concepto está próximo al razonamiento clínico y la resolución de problemas. El siguiente nivel “*mostrar como*” refleja las habilidades del alumno para actuar adecuadamente en una situación práctica y describe habilidades técnicas y de comportamiento en una situación práctica o simulada. El vértice de la pirámide es el nivel “*hacer*” que se refiere a la actuación real en la práctica clínica. Los 4 niveles de competencia de Miller y la asociación de las evaluaciones que hace son probablemente el mejor sistema de

estratificación conocido para describir la medición del aprendizaje en educación médica. La mayoría de revisiones sobre la evaluación de las competencias se han basado en este modelo [88].

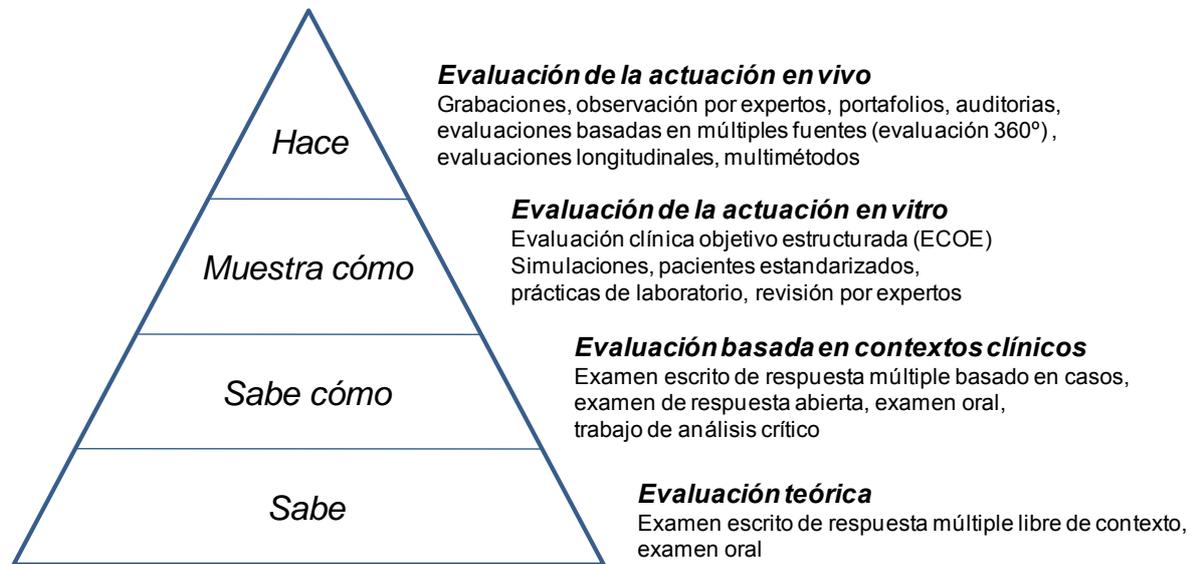


Figura 2. Pirámide de la competencia de Miller con ejemplos de técnicas de evaluación utilizadas en educación médica.

La opinión de los expertos respecto a la evaluación de las competencias profesionales en educación médica es que se han de evaluar no solo la memoria, el reconocimiento de determinados hechos y la demostración de las habilidades prácticas sino también la capacidad de los alumnos de sintetizar la información en un contexto determinado y su capacidad de aplicarla en aquellas situaciones que precisan del pensamiento crítico y la resolución de los problemas. El objetivo último de la evaluación en educación médica es determinar la capacidad de los alumnos de integrar e implementar los diferentes dominios o habilidades del aprendizaje que definen la competencia profesional, de forma consistente a lo largo del tiempo en un entorno lo más próximo posible al real [104].

El problema es que todavía no se ha diseñado un único método de evaluación ni un único formato válido para medir todos los niveles de la competencia clínica. Para cada nivel de competencia tampoco existe una evaluación estandarizada y validada. Cuanto más alto es el nivel de competencia que mide la pirámide más autenticidad clínica ha de tener la evaluación. Son diversas las razones por las que es difícil evaluar la competencia profesional: la dificultad de medir el conocimiento tácito, la imposibilidad de formular preguntas clínicas específicas para cada competencia que permitan su medición, la falta de acuerdo sobre cuando se considera que una competencia es aceptable y por las múltiples influencias que pueden intervenir [107].

Debido a la complejidad de la competencia clínica, probablemente deberían usarse evaluaciones múltiples y diferentes a lo largo del tiempo [108]. Según el método de evaluación que se emplee los resultados al comparar distintos métodos de enseñanza pueden diferir porque la habilidad o competencia que estamos midiendo será distinta. Hemos de asegurar, por consiguiente, que las

evaluaciones sean apropiadas para las habilidades que se estén midiendo [88,99,104]. Otro problema es que dispongamos de instrumentos de evaluación para todas las competencias. Es necesario explorar nuevos métodos de evaluación que capturen la complejidad de las intervenciones, que se basen en los resultados esperados y que sean fiables, válidos, flexibles, comprensibles, viables, que midan las competencias en tiempo real y eficaces [95,99].

Los exámenes escritos, orales y los tests de respuesta múltiple son adecuados para medir el conocimiento básico (*“saber”*). Los tests de respuesta múltiple tienen la gran ventaja de su fiabilidad pero tienen una validez limitada: ofreciendo opciones se está condicionando la respuesta de los alumnos y desaparece la posibilidad de que el conocimiento se genere activamente [88]. Los tests de respuesta múltiples son muy útiles para medir el conocimiento explícito (memoria o reconocimiento), que es el nivel de conocimiento más básico, pero no otros niveles del conocimiento más complejos, como la comprensión o la aplicación del conocimiento a la resolución de problemas [96,109]. Tampoco son válidos para predecir las competencias clínicas [99]. Los exámenes de respuesta abierta y los exámenes orales son más adecuados para medir la comprensión [96], un nivel de función cognitiva superior a la memoria o reconocimiento [109]. Los exámenes orales también presentan una fiabilidad aceptable [110] pero también tienen limitada su validez [99].

Todos estos formatos requieren ser adaptados para evaluar el conocimiento aplicado a diferentes contextos, la toma de decisiones, la resolución de problemas o el razonamiento clínico (*“saber como”*). Para ello, se recomienda incluir escenarios clínicos o presentación de casos en la evaluación. Una limitación potencial de este tipo de evaluaciones es su escasa fiabilidad y capacidad de repetición; para mejorar este aspecto se recomienda que el formato de preguntas sea lo más estructurado posible y que el diseño de la evaluaciones delimite con claridad las áreas de las competencias que se quieren analizar [88]. Distinguir entre las preguntas que analizan la aplicación del conocimiento y las que reflejan la mera adquisición del mismo ayuda en este sentido, sin embargo, este supuesto no es fácil de conseguir [111]. Medir sobre una plantilla los elementos que presenta el caso también ayuda a aumentar su fiabilidad [112].

Para evaluar el tercer nivel de las competencias (*“mostrar como”*) son necesarios las evaluaciones clínicas objetivo estructuradas (ECO) [101,113], las evaluaciones basadas en pacientes estandarizados, las simulaciones, las prácticas en el laboratorio o las revisiones por expertos. En Anestesiología, como en la mayoría de campos, ha sido limitado el entusiasmo generado por las ECO. Hay serios problemas de coste, viabilidad y fiabilidad y faltan publicaciones que evidencien su eficacia [99]. Estas evaluaciones no son apropiadas para valorar aspectos de la competencia profesional como las relaciones interpersonales, las actitudes o las consideraciones éticas y el profesionalismo [114]. La simulación ha demostrado una eficacia considerable como método de evaluación en Anestesiología, sobre todo de la actuación en relación a los incidentes críticos. Los simuladores son buenas herramientas de evaluación tanto de las habilidades técnicas [115] como de las no técnicas [116] sobre todo cuando se asocian a grabaciones en video. Respecto al paciente estandarizado como sistema de evaluación es poco viable en Anestesiología. La mayoría de situaciones clínicas relevantes en Anestesiología no son fácilmente trasladables a este tipo de

evaluación (excepto para el paciente con dolor crónico); además a diferencia de las especialidades médicas, en Anestesiología no es fácil reclutar estos pacientes [99].

Por último, evaluar el vértice de la *pirámide de Miller* (“*hacer*”) es un reto para todos los implicados en la evaluación de las competencias clínicas. Este nivel enfoca la evaluación de la competencia profesional en global (competencia general), no en sus componentes aislados. Las grabaciones de la actuación en situaciones reales, la observación por expertos, los cuadernos de autoaprendizaje o portafolios, la combinación de múltiples métodos de evaluación, las evaluaciones longitudinales a lo largo del tiempo, las auditorias y las evaluaciones basadas en múltiples fuentes (evaluaciones 360°) tienen importantes limitaciones logísticas. El resultado clínico del paciente no es un marcador adecuado de la competencia profesional porque está sujeto a la influencia de muchos factores no relacionados directamente con la actuación del profesional [107,114].

Los sistemas de evaluación de la competencia profesional basados en la observación por anesestiólogos expertos [107] son métodos viables, sobre todo si incorporan esquemas de valoración, pero de baja fiabilidad ya que son subjetivos; la fiabilidad puede mejorar si las evaluaciones las hacen varios evaluadores diferentes [117,118], el observador no está involucrado en el cuidado del paciente [99] o se incorporan herramientas como la grabación de las actuaciones [114]. Este método de evaluación fue validado indirectamente por Rhoton et al. [119] quienes observaron una correlación entre las deficiencias observadas en las habilidades no cognitivas (confianza, compostura, habilidades interpersonales, deseo de aprender, comportamiento profesional) y los incidentes críticos. Greaves y Grant [107] también encuentran patrones inapropiados de comportamiento asociados a la incompetencia; estos incluyen falta de flexibilidad, fracaso en adaptarse a los cambios en las circunstancias, juicios irracionales y actuaciones desordenadas. Ambos estudios evidencian la dificultad de medir la competencia profesional entendida como un todo [104,108].

El portafolio es un registro de aprendizaje que hace el propio alumno con el objetivo de evidenciar su progreso. Los alumnos presentan su trabajo a lo largo del tiempo utilizando metodologías de trabajo diferentes: imágenes, registros de pacientes, casos clínicos, proyectos, pósters y abstracts, copia de evaluaciones, así como comentarios de autovaloración y reflexión sobre su proceso de aprendizaje. La revisión periódica de los portafolios junto al docente permite determinar los logros y progresos del alumno en la adquisición de competencias y planificar los objetivos de aprendizaje [96]. A pesar de sus limitaciones logísticas el portafolio parece tener potencial en el futuro como sistema de evaluación de la competencia profesional en Anestesiología [99]. Respecto a las evaluaciones basadas en múltiples fuentes es un sistema de evaluación integral que se basa en el equilibrio o triangulación entre las evaluaciones que miden el conocimiento en su nivel de aplicación o de resolución de problemas [109], las habilidades adquiridas a lo largo del tiempo y el registro del juicio auto-crítico y la reflexión del alumno [120,121].

En una excelente revisión publicada en 2007 en el *New England Journal of Medicine* [122], Epstein analiza el impacto, los efectos no intencionados y las limitaciones de los distintos métodos de evaluación en educación médica. El autor concluye afirmando que el contenido, el formato y la

frecuencia de las evaluaciones, así como el momento y el formato de su potencial uso formativo han de adaptarse al entorno y a los objetivos específicos de cada programa de formación. La revista *Medical Teacher* también tiene publicadas Guías sobre distintos métodos de evaluación de las competencias en Educación Médica Avanzada de la *Association for Medical Education in Europe* [123-126]. Diseñar métodos de evaluación fiables que puedan predecir con validez las competencias profesionales en la práctica clínica y que al mismo tiempo tengan un papel educativo es un *gold estándar* que todavía no se ha conseguido.

La adquisición de conocimiento, la capacidad de aplicar el conocimiento para la solución de problemas, el manejo de la información (habilidades cognitivas) y usar el razonamiento clínico o unir el conocimiento básico y clínico (habilidades integrativas) [104], (Tabla 4), son habilidades comunes a los tres niveles de formación en Anestesiología: *pregrado*, *postgrado* y *Formación Médica Continuada*; nosotros enfocamos nuestra investigación en el análisis de estas habilidades. El diseño de nuestro estudio ha enfocado el análisis de la eficacia de la enseñanza en el *Nivel 2B* del modelo de evaluación de resultados modificado de Kirkpatrick (Tabla 3) y en el nivel “*Saber cómo*” (*conocimiento aplicado*) de la pirámide de evaluación de competencias de Miller (Figura 2); con tal fin, hemos diseñado evaluaciones que se ajusten a los objetivos docentes fijados y basadas en el contexto clínico. El formato de *evaluación basada en casos o situaciones clínicas* [127] es una alternativa válida y fiable para analizar los resultados de la enseñanza y las competencias de los alumnos en estos niveles de evaluación [104,111].

2.6. Justificación del estudio

Los programas de formación tienen el reto de adaptarse a las exigencias de los cambios que se producen en su entorno: cambios de la sociedad de la información que obligan a formar a los alumnos en el manejo de una gran cantidad de información; un perfil de los alumnos más heterogéneo; cambios en los avances científicos que requieren una actitud crítica y abierta; cambios derivados de la incorporación de las tecnologías de la información y comunicación; una sociedad más compleja y diversa que necesita alumnos que sepan enfrentarse a situaciones cambiantes y complejas o el proceso de adaptación al *Espacio Europeo de Educación Superior*. Por este motivo, las instituciones han de adaptar los procesos de enseñanza buscando acentuar la participación activa del alumno en su proceso de aprendizaje, enfatizando la adquisición de habilidades más allá de la pura adquisición de conocimientos y otorgándoles herramientas que permitan enfrentarse al proceso de formación continuada. En este contexto, parece indiscutible la necesidad de introducir nuevos métodos docentes e investigar si esta innovación mejora la eficacia de la enseñanza [128]. Los rápidos cambios que acontecen en la investigación, en la práctica clínica y en las tecnologías de la información justifican la necesidad de habilidades de aprendizaje más duraderas en educación médica, y por consiguiente la necesidad de desarrollar nuevos métodos o estrategias de enseñanza de estas habilidades [129].

Los programas de formación en Anestesiología han de adoptar las medidas que sean más eficaces para cumplimentar las necesidades docentes de los alumnos. Esto solo será posible si se sistematiza la investigación docente en este campo de la medicina y ésta sirve de base para estudiar y comparar

los distintos métodos de enseñanza. En Anestesiología, en particular, son escasos los estudios que se han realizado en educación médica. Los aspectos metodológicos de la enseñanza de la Anestesiología son un campo de investigación poco explorado. Hasta la fecha, no hemos encontrado en la literatura estudios prospectivos que comparen la eficacia del *aprendizaje basado en la discusión casos/problema* con la *clase magistral* en el campo de la Anestesiología. En otros campos de la educación médica los resultados de estudios similares han sido contradictorios [49,50,76,130,131].

Las barreras potenciales a esa comparación han podido incluir las preferencias del profesorado por la *clase magistral*, la falta de conocimiento y difusión del método *aprendizaje basado en la discusión de casos/problema*, la necesidad de formación y entrenamiento del profesorado en este método o la necesidad de adaptación de los contenidos docentes al nuevo método. En nuestra opinión, estas barreras no son infranqueables. Revisiones recientes sugieren la necesidad de investigar la eficacia de la enseñanza en educación médica diseñando estudios metodológicamente más rigurosos, que comparen metodologías que estimulen la reflexión y el aprendizaje, que definan con claridad el nivel de resultados y competencia que quieren medir, que incluyan medidas de calidad de las evaluaciones, que intenten correlacionar diferentes mediciones y valoren los resultados en el tiempo [95,132].

En la formación médica de *pregrado*, la enseñanza del *soporte vital básico* es uno de los temas principales del programa de formación de los estudiantes de medicina. La razón es que este tema tiene una trascendencia socio-sanitaria más allá del ámbito de la especialidad. Los objetivos primarios de la enseñanza del *soporte vital básico* son aprender la secuencia de actuación según las recomendaciones de las Guías internacionales y adquirir las habilidades necesarias para su correcta aplicación [133,134]. A pesar de los esfuerzos de preparación y actualización de las Guías de *soporte vital básico* todavía no existen recomendaciones sobre cuál es el mejor método de enseñanza [135,136]; recientemente se ha propuesto algún modelo para la enseñanza de profesionales no sanitarios [137]. Para establecer niveles significativos de evidencia, el *International Liaison Committee on Resuscitation (ILCOR)* ha adoptado como objetivo el diseño de estudios randomizados y controlados en investigación docente que comparen los métodos de enseñanza tradicionales basados en el profesor con métodos más novedosos enfocados en el proceso de participación de los alumnos [138]. Siguiendo las recomendaciones de la *ILCOR*, hemos incluido el *aprendizaje basado en la discusión de casos/problema* en la enseñanza del *soporte vital básico* y hemos enfocado la evaluación en los aspectos relacionados a la aplicación del conocimiento.

La *valoración pre-anestésica* es un tema común en todos los programas de *formación de residentes de Anestesiología* de primer año [páginas web: <http://www.academia.cat/societats/dolor/fmc.htm>; <http://www.rcoa.ac.uk/docs/CCTpti.pdf>]. Es un proceso complejo en el que la información debe ser recogida, agrupada y analizada. Esta información puede utilizarse para organizar los recursos perioperatorios, establecer el plan anestésico, indicar los cuidados postoperatorios y el establecer estrategias para el control del dolor. Como este tema va más allá de la mera adquisición de conocimientos teóricos, ofrece un contexto ideal para la valoración de la eficacia de nuevos métodos de enseñanza.

El *embolismo aéreo* en anestesia es una complicación potencialmente fatal donde el anesestesiólogo juega un papel crucial en reducir la morbi-mortalidad gracias a un diagnóstico y tratamiento precoces [139-141]; por este motivo, este es un tema esencial en los programas de *Formación Médica Continuada* en Anestesiología. Desde 1986, la *Fundación Europea para la Enseñanza en Anestesiología (FEEA)*, actualmente denominado *Comité para la Educación Europea en Anestesiología (CEEA)*, viene desarrollando un programa de *Formación Médica Continuada* de amplia aceptación [142]. Cada curso organizado por el centro *CEEA* en Cataluña, incluye 30 horas lectivas que se imparten en 2,5 días. El método docente más frecuentemente utilizado es la *clase magistral*, que es como tradicionalmente se enseña el tema del *embolismo aéreo* en anestesia, pero el método *aprendizaje basado en la discusión de casos/problema* puede ser también un formato adecuado [143].

Los responsables de la formación en Anestesiología debemos de analizar la eficacia de los métodos de enseñanza para fundamentar su elección y utilización, así como conocer las particularidades de las distintas técnicas de evaluación. La docencia y su evaluación han de adaptarse a las necesidades de aprendizaje de los alumnos y a los recursos disponibles. Los educadores hemos de saber ponderar estos factores para tomar las decisiones más adecuadas. La educación en Anestesiología está especialmente necesitada de investigaciones que proporcionen esta información tanto a los profesores como a los alumnos.

Tres son las razones principales que nos han conducido a la presente investigación: 1) la falta de evidencia objetiva en la literatura sobre la eficacia de los métodos de enseñanza de la Anestesiología [129]; 2) las ventajas potenciales de un método novedoso, el *aprendizaje basado en la discusión de casos/problema*, que hemos incorporado a nuestros programas docentes y los datos prometedores en la literatura de su aplicación en Anestesiología y, sobre todo, 3) el estímulo de los alumnos y el entusiasmo del equipo docente por la investigación en este campo de la Medicina.

En base a estas razones y en concordancia con las recomendaciones para la investigación en educación médica [83] hemos diseñado el presente estudio comparando los dos métodos que empleamos en la enseñanza de la Anestesiología: el método tradicional basado en la *clase magistral* y el método más nuevo de *aprendizaje basado en la discusión de casos/problema*. Hemos empleado un diseño experimental en el contexto de tres programas de formación en Anestesiología que incluye contenidos específicos para cada nivel de formación. Hemos enfocado la evaluación en aspectos relacionados con la aplicación práctica del conocimiento (Figura 3). Creemos que el diseño metodológico de esta investigación puede proporcionar evidencia sobre la eficacia de los métodos de enseñanza en Anestesiología.

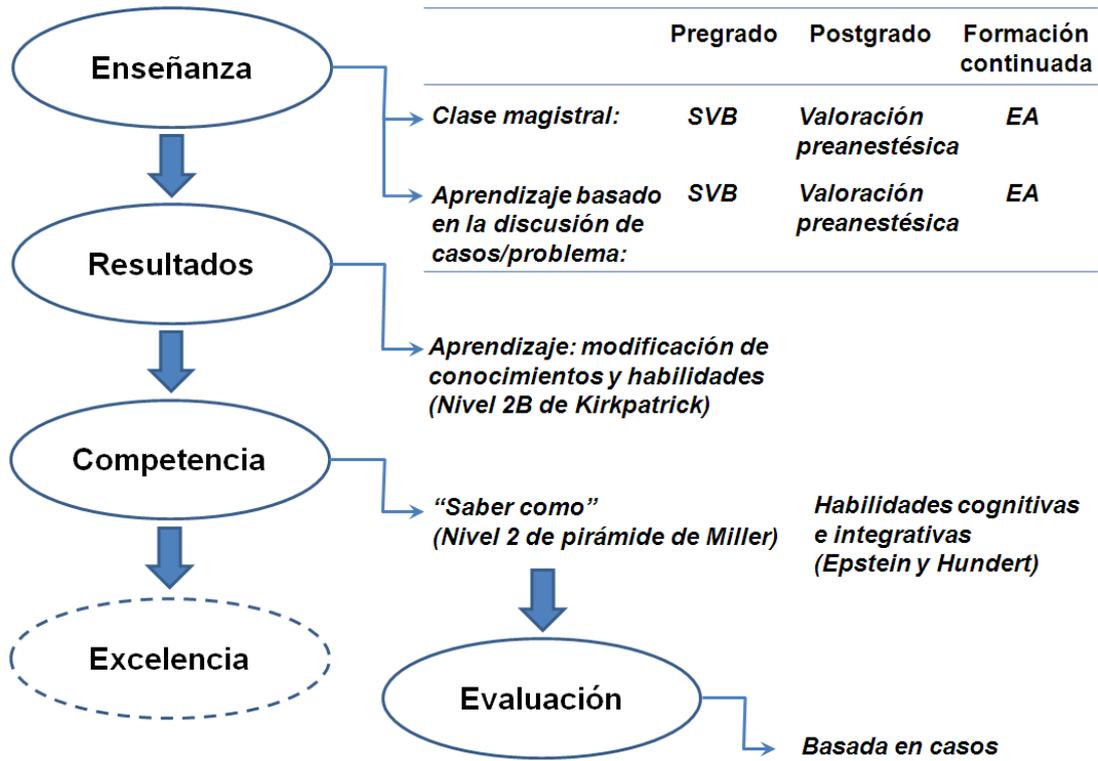


Figura 3. Componentes del diseño de investigación: métodos docentes, contexto y temas de enseñanza; nivel de resultados, competencias analizadas y método de evaluación aplicado. SVB: soporte vital básico; EA: embolismo aéreo.

3. HIPOTESIS Y OBJETIVOS

Hipótesis

El método *aprendizaje basado en la discusión de casos/problema* es más eficaz que el método tradicional basado en la *clase magistral* en la adquisición de conocimientos inmediatos y en mejorar las habilidades cognitivas e integrativas de los alumnos. Como cualquier otro método de enseñanza, su eficacia ha de ser independiente de la materia que se explica y su aplicabilidad demostrarse en *pregrado, postgrado y Formación Médica Continuada*.

Objetivo general

1. Valorar y comparar la eficacia del *aprendizaje basado en la discusión de casos/problema* con el método de enseñanza tradicional basado en la *clase magistral* en la enseñanza de la Anestesiología.

Objetivos específicos

1. Valorar si el método de enseñanza basado en la *discusión de casos/escenario* simulados en maniqués por el profesor mejora la capacidad de los estudiantes de medicina de aplicar el conocimiento adquirido sobre *soprote vital básico* comparado con el método tradicional de enseñanza basado en la *presentación multimedia*.

Artículo 1. **Carrero E**, Gomar C, Penzo W, Fabregas N, Valero R, Sanchez-Etayo G. Teaching basic life support algorithms by either multimedia presentations or case based discussion equally improves the level of cognitive skills of undergraduate medical students. *Med Teach* 2009; Feb 24:e1-e7. Disponible en: URL; <http://dx.doi.org/10.1080/01421590802512896>

2. Valorar y comparar la eficacia del método *aprendizaje basado en la discusión de casos/problema* y el método tradicional de *clase magistral* en la enseñanza del tema “*valoración pre-anestésica*”, incluido en el programa de formación teórica de los residentes de Anestesiología de primer año de la Societat Catalana d’Anestesiología, Reanimació i Terapèutica del Dolor, utilizando una herramienta objetiva para la evaluación del conocimiento antes y después de la enseñanza.

Artículo 2. **Carrero E**, Gomar C, Penzo W, Rull M. Comparison between lecture-based approach and case/problem-based learning discussion for teaching pre-anaesthetic assessment. *Eur J Anaesthesiol* 2007; 24:1008-15.

3. Comparar la eficacia de la *clase magistral* y el *aprendizaje basado en la discusión de casos/problema* en la adquisición de conocimientos sobre *embolismo aéreo* en anestesia como parte del programa de *Formación Médica Continuada* del centro del Comité para la Educación Europea en Anestesiología en Cataluña.

Artículo 3. **Carrero E**, Gomar C, Fábregas N, Penzo W, Castillo J, Villalonga A. Clase magistral versus aprendizaje basado en caso/problema para la enseñanza del embolismo aéreo en formación médica continuada. *Rev Esp Anesthesiol Reanim* 2008; 55: 202-9.

4. MATERIAL Y MÉTODOS

4.1. Metodología general

4.1.1. Diseño

Los tres artículos de la presente tesis forman parte de una misma línea de investigación y presentan una metodología común que comparten y otra específica. Cada artículo está enmarcado en un nivel de formación determinado (*pregrado, postgrado y formación continuada*). El diseño del estudio lo adaptamos a los tres programas de formación ya establecidos para que las modificaciones del programa fuesen mínimas, respetando siempre el tiempo programado de las clases.

Los tres estudios fueron controlados, prospectivos y randomizados. Comparamos dos grupos en virtud del método de enseñanza aplicado. El *grupo control* incluyó a los alumnos que recibían la enseñanza del tema seleccionado con el método tradicional basado en la *clase magistral* (Apartado 4.1.2). El *grupo problema* incluyó a los alumnos a los que se aplicó el método nuevo basado en la *discusión de casos/problema* (Apartado 4.1.3).

En primer lugar los alumnos firmaron el consentimiento informado de su participación en el estudio. Para garantizar que los grupos fueran comparables entre sí, los alumnos cumplieron una de las evaluaciones antes de la aplicación del método de enseñanza (*evaluación pre*) (Apartado 4.1.4).

La randomización se realizó inmediatamente después de la *evaluación pre*. Se distribuyeron sobres cerrados a los alumnos con los números del 1 al número máximo de alumnos de cada curso que se incluyeron en la randomización. Se excluyeron de la misma los alumnos que no firmaron el consentimiento informado o no completaron la *evaluación pre*. Los alumnos quedaron randomizados por programa informático, restringido a dos bloques, para recibir una *clase magistral* (*grupo clase magistral*) o una sesión de *discusión de casos/problema* (*grupo caso/problema*). Esta aleatorización selectiva se realizó para garantizar que el número de sujetos en cada grupo fuese el mismo y optimizar así el análisis de los resultados.

Las características de este tipo de estudios hacen imposible que los alumnos y los profesores sean ciegos al método de enseñanza aplicado. Los evaluadores que puntuaron las evaluaciones sí fueron ciegos a la asignación de los grupos.

El diseño del estudio se presentó para su aprobación por la Comisión de Investigación del Departamento de Cirugía y Especialidades Quirúrgicas de la Facultat de Medicina de la Universitat Central de Barcelona, a la Comisión de Docencia de la Societat Catalana d'Anestesiologia, Reanimació i Terapèutica del Dolor y a la Dirección del Comité para la Educación Europea en Anestesiología, Centro de Cataluña.

4.1.2. Sujetos

4.1.2.1. Cálculo del tamaño de la muestra

Se aplicó la fórmula del cálculo del tamaño de la muestra para contraste de hipótesis, comparación de proporciones, grupos independientes e hipótesis bilateral [144]:

$$n = \frac{\left[Z_{\alpha} \cdot \sqrt{2p(1-p)} + Z_{\beta} \cdot \sqrt{p_1(1-p_1) + p_2(1-p_2)} \right]^2}{(p_1 - p_2)^2}$$

Donde:

- n = alumnos necesarios en cada uno de los grupos
- Z_{α} = valor Z correspondiente al riesgo α fijado
- Z_{β} = valor Z correspondiente al riesgo β fijado
- p_1 = valor de la proporción en el grupo control (*grupo clase magistral*)
- p_2 = valor de la proporción en el grupo problema (*grupo caso/problema*)
- p = media aritmética de las dos proporciones p_1 y p_2

Se estimó en los tres artículos un riesgo $\alpha = 0.05$ (nivel de confianza del 95%) y un riesgo $\beta = 0.2$ (potencia estadística del 80%); o lo que es lo mismo, un 5% de probabilidades de no aceptar la hipótesis nula siendo esta cierta y un 20% de probabilidad de aceptar la hipótesis nula siendo esta falsa. Los valores α y β se decidieron de acuerdo a los valores utilizados habitualmente en la literatura. Aunque la hipótesis de nuestro estudio era unilateral (demostrar si el método *caso/problema* es mejor que la *clase magistral*) aplicamos el valor de Z_{α} para hipótesis bilateral porque es la estimación más conservadora. Los valores p_1 y p_2 se estimaron en base a las publicaciones relacionadas con la eficacia media de la enseñanza de la *clase magistral* y de la *discusión de casos problema* en *pregrado*, *postgrado* y *formación continuada*. Consideramos relevante una diferencia de puntuación ≥ 25 ó 30% dependiendo del curso de formación.

Por último, ajustamos el tamaño de la muestra a las pérdidas estimadas de casos en cada curso según la fórmula:

$$\text{Muestra ajustada a las pérdidas} = n (1 / 1 - R)$$

Donde: n = muestra calculada y R = proporción calculada de pérdidas

4.1.2.2. Población del estudio

El estudio se desarrolló en alumnos matriculados de tres programas oficiales de tres niveles diferentes de formación de Anestesiología. La población del estudio incluyó alumnos de cada uno de los niveles de formación:

1. *Pregrado*: programa de formación en *soporte vital básico* del Departamento de Cirugía y Especialidades Quirúrgicas de la Facultat de Medicina de la Universitat de Barcelona. Alumnos de medicina de tercer año en la asignatura Fundamentos de Cirugía. El material docente de la Facultad de Medicina de la UB se encuentra en la Web: <http://dossiers.ub.edu/assi.php?a=1834>
2. *Postgrado*: programa de formación teórica para residentes de Anestesiología de primer año de la Societat Catalana d'Anestesiologia, Reanimació i Terapèutica del Dolor (SCARTD). Residentes de

Anestesiología de primer año. El material docente de la SCARTD se encuentra en la página Web: www.academia.cat/societats/dolor/fmc.htm

3. *Formación continuada*: programa de *Formación Médica Continuada* del Comité para la Educación Europea en Anestesiología (CEEA), centro de Cataluña [142]. Anestesiólogos de distintas comunidades autónomas de España. El material docente del CEEA se encuentra en la página Web: www.euroviane.net

La participación en el estudio fue voluntaria, anónima y no tuvo influencia en la puntuación. A todos los participantes se les entregó una hoja con las instrucciones para cumplimentar las evaluaciones y el consentimiento informado (Tabla 5).

Tabla 5. Hoja de instrucciones para cumplimentar las evaluaciones y el consentimiento informado.

Instrucciones

1. *Pregrado*: A continuación se le expondrán tres preguntas de respuesta abierta en formato de casos clínicos. Es importante que su respuesta sea lo más específica posible de acuerdo a la secuencia de soporte vital básico. Posteriormente se le mostrará una grabación de video que muestra a un solo reanimador ejecutando la secuencia de soporte vital básico. La secuencia de soporte vital básico emitida puede contener maniobras realizadas incorrectamente o no. Por favor, indique las maniobras incorrectas, si las hubiere, que detecte.
Postgrado: A continuación tiene expuesto dos casos clínicos sobre los que se formulan distintas preguntas. Observará que los datos que incluye el caso están espaciados para su correcta identificación y revisión posterior. Es importante que razone sus respuestas sólo cuando lo solicite la pregunta.
Formación continuada: A continuación tiene expuesto un caso clínico sobre el que se formulan distintas preguntas. Observará que los datos que incluye el caso y los espacios para las respuestas se han dispuesto en forma vertical con la finalidad de facilitar su revisión posterior. No debe justificarse necesariamente cada dato ni rellenarse todas las casillas previstas sino sólo aquellas que, bajo su criterio profesional, sirven para contestar la pregunta formulada.
2. Es muy importante que conteste de forma individual.
3. Los resultados serán analizados de forma global, sin valoración individual.
4. El tiempo máximo para rellenar la evaluación es de 15 minutos (20 minutos en *postgrado*). Se avisará 5 minutos antes. Es importante ser estrictos en el tiempo.
5. Para garantizar el anonimato a la hora de procesar los datos registrados se adjudicará aleatoriamente un sobre cerrado con un número personal en su interior; este número debe figurar en el encabezado de las hojas de evaluación. Es importante recordar el número adjudicado y guardar la tarjeta hasta el final de la estación de trabajo, clase o curso.
6. Al final de la estación de trabajo, clase o curso, para quien lo desee, se discutirá cualquier cuestión sobre el caso expuesto.

Consentimiento informado

1. Doy mi consentimiento a participar en el presente estudio de investigación sobre adquisición de conocimientos y habilidades en Anestesiología.
2. Mi participación es voluntaria, siendo libre de abandonar en cualquier momento el estudio.
3. Nombre y Apellidos:
4. D.N.I y Firma:

4.1.2.3. Criterios de inclusión

Los criterios de inclusión del estudio fueron:

1. Firma del consentimiento para participar en el estudio.
2. Asistencia a la clase.
3. Cumplimentación de las *evaluaciones pre y post*.

Se excluyeron aquellos participantes que no cumplieron todas las fases del estudio. En cada artículo se han incluido una figura que muestra el diagrama de randomización y participación (diagrama de flujo) y otra figura de las etapas del estudio.

4.1.3. Enseñanza

Los estudios se realizaron sobre una única estación de trabajo, clase o sesión docente. El lugar para la enseñanza fue el mismo que estaba pre-establecido antes del estudio en cada programa de formación. Los contenidos docentes fueron definidos con detalle y consensuados con anterioridad a cada estudio por el Comité Docente del programa de acuerdo a las guías de evidencia científica publicadas sobre cada tema. Dos profesores participaron en cada estudio; los profesores tenían experiencia docente en el tema y se reunieron antes de los estudios para ponerse de acuerdo sobre los procedimientos a aplicar. Los objetivos docentes de cada uno de los tres artículos fueron los mismos en ambos grupos. El tiempo total de la enseñanza fue el mismo con los dos métodos docentes.

Se mantuvo, en cada ámbito del estudio, el formato original de *clase magistral*: exposición del profesor con soporte audiovisual seguido de preguntas.

En el *grupo caso/problema* se empleó todo el tiempo de la clase para la discusión de los casos. El entorno de la clase se adaptó, distribución en semicírculo, para facilitar la participación de los alumnos [10] y el profesor trató de dirigirse a los alumnos por su nombre para favorecer la discusión [58]. En *pregrado* el planteamiento del caso se escenificó sobre un maniquí de *soporte vital básico*; en *postgrado* y *formación continuada* se plantearon casos clínicos.

La *discusión de los casos/problema* se presentó como una sesión única y los alumnos no recibieron los casos por adelantado. El profesor estimulaba la participación de los alumnos planteando preguntas relacionadas con los objetivos docentes. La discusión la dirigió y moduló el profesor que detectaba y corregía los errores y dudas de los alumnos. El profesor era experto en el tema, en la metodología docente y tenía experiencia clínica en los casos planteados.

4.1.4. Evaluaciones y estudio piloto

Las evaluaciones midieron habilidades cognitivas e integrativas [104]. El formato de las evaluaciones se basó en la presentación de situaciones clínicas hipotéticas que cumplieran los objetivos docentes.

Previamente a cada una de las tres partes del estudio se realizó una *prueba piloto*. Los estudiantes (*pregrado*), residentes (*postgrado*) y anestesiólogos (*formación continuada*) que la cumplieron no participaron en el estudio. Los evaluadores que puntuaron la prueba piloto fueron los mismos que posteriormente puntuarían a los participantes en el estudio. Se confirmó que las evaluaciones eran comprensibles, del mismo grado de dificultad y fáciles de completar en 15 minutos (artículos 1 y 3) ó

20 minutos (artículo 2). Esta prueba piloto sirvió para eliminar de las evaluaciones los *ítems* o preguntas que causaban discrepancia entre los profesores y así generar el diseño final de las evaluaciones.

El estudio incluyó dos evaluaciones, una antes de la enseñanza (*evaluación pre*) y otra inmediatamente después de la enseñanza (*evaluación post*). Las *evaluaciones pre* y *post* fueron la misma en el estudio de *pregrado* y equivalentes en los estudios de *post-grado* y *formación continuada*. El primer día, se decidió aleatoriamente (artículos 2 y 3) que evaluación cumplimentaría cada grupo antes de la enseñanza y cual después de la enseñanza. Los participantes no pudieron realizar comentarios o preguntas relacionados con las evaluaciones hasta que finalizó la *evaluación post*.

Las variables medidas en las evaluaciones fueron diferentes campos del conocimiento. Cada campo incluía un número determinado de *ítems* sobre los que se contabilizó el número total de respuestas correctas. Los *ítems* contestados correctamente puntuaron 1 punto y los *ítems* no contestados o contestados incorrectamente puntuaron 0 puntos. Se calculó la puntuación obtenida para cada uno de los campos en la *evaluaciones pre* y *post*. También se calculó la diferencia de puntuaciones entre las *evaluaciones post* y *pre*. Por último, se analizaron si existían diferencias entre grupos e intragrupo.

Consideramos apropiado el análisis de la correlación lineal entre los campos de conocimiento con rangos de puntuación altos (artículo 2: campo 1 y campo 2; artículo 3: *ítems* 1 y 2 del campo 1). Estudiamos en cada grupo si existía correlación lineal entre el reconocimiento (habilidad cognitiva) y el razonamiento (habilidad integrativa) en las evaluaciones antes y después de la enseñanza.

Las evaluaciones fueron puntuadas por dos evaluadores diferentes a los dos profesores que impartieron la enseñanza, ciegos a la asignación de grupos. Empleamos el estadístico “*kappa*” para analizar la concordancia entre evaluadores [145].

4.1.5. Análisis estadístico

Los cálculos estadísticos se realizaron usando el paquete estadístico SSPS 10.0 ó 13.0 (SSPS Inc, Chicago IL, USA). Las variables cualitativas se expresaron en números absolutos (proporciones) y se compararon usando el test Chi-cuadrado. El test exacto de Fisher se empleó para las tablas de frecuencias 2x2 cuando los valores de las frecuencias esperadas fueron inferiores a 5. Las variables cuantitativas no siguieron una distribución normal. Los datos se expresaron como medianas y percentiles 25-75. Se utilizaron pruebas no paramétricas para analizar las diferencias entre grupos (test de la U de Mann-Whitney para dos muestras independientes, test de Kruskal Wallis para *k* muestras independientes) e intragrupo (test de Wilcoxon para dos muestras apareadas o test de McNemar para variables dicotómicas y test de Friedman para *k* muestras apareadas o test de la Q de Cochran para variables dicotómicas). Calculamos el coeficiente de correlación de Spearman (*r*) y el coeficiente de determinación (r^2) para determinar si existía correlación lineal entre dos variables y el grado de la correlación. Definimos la significación estadística como $p < 0,05$.

4.2. Metodología Artículo 1. Carrero E, Gomar C, Penzo W, Fabregas N, Valero R, Sanchez-Etayo G. Teaching basic life support algorithms by either multimedia presentations or case based discussion equally improves the level of cognitive skills of undergraduate medical students. *Med Teach* 2009; Feb 24:e1-e7. Disponible en: URL; <http://dx.doi.org/10.1080/01421590802512896>

4.2.1. Diseño

El diseño del estudio en *pregrado* se adaptó al programa de formación en *soporte vital básico* de los estudiantes de medicina. En la Facultad de Medicina de la Universitat de Barcelona, la enseñanza del *soporte vital básico* está incluida en el currículo como una estación de trabajo teórico-práctica de 2,5 horas dirigida por el profesorado. Los algoritmos de actuación se enseñan durante la primera hora tradicionalmente mediante una grabación de video doméstica más una presentación Power Point; durante la siguiente hora y media los estudiantes practican sobre maniqués la secuencia de acción del *soporte vital básico*, el manejo de la vía aérea, la ventilación, las compresiones torácicas externas y la desfibrilación semi-automática. La evaluación de los estudiantes es teórica y está incluida como preguntas tipo test de respuesta múltiple en el examen final del curso académico.

El diseño de esta parte del estudio cumplió las características de la metodología general del estudio (Apartado 4.1.1.).

4.2.2. Sujetos

4.2.2.1. Cálculo del tamaño de la muestra

Se estimó una eficacia media para la *clase magistral* del 60% ($p_1=0,6$); consideramos relevante una diferencia entre las puntuaciones $\geq 30\%$ ($p_2 = 0,9$) y estimamos una pérdida de casos aproximada del 5% ($R=0,05$). El tamaño de muestra resultante fue 33 estudiantes en cada grupo.

4.2.2.2. Población del estudio

La población del estudio incluyó 70 alumnos de tercer año de la Facultat de Medicina de la Universitat de Barcelona de la asignatura Fundamentos de Cirugía, que participaron en cuatro estaciones consecutivas de formación en *soporte vital básico*. En esta población registramos la experiencia y formación previa en reanimación cardiopulmonar de los alumnos. La Figura 4 muestra el diagrama de randomización y participación de los alumnos en el estudio y la Figura 5 las etapas del estudio.

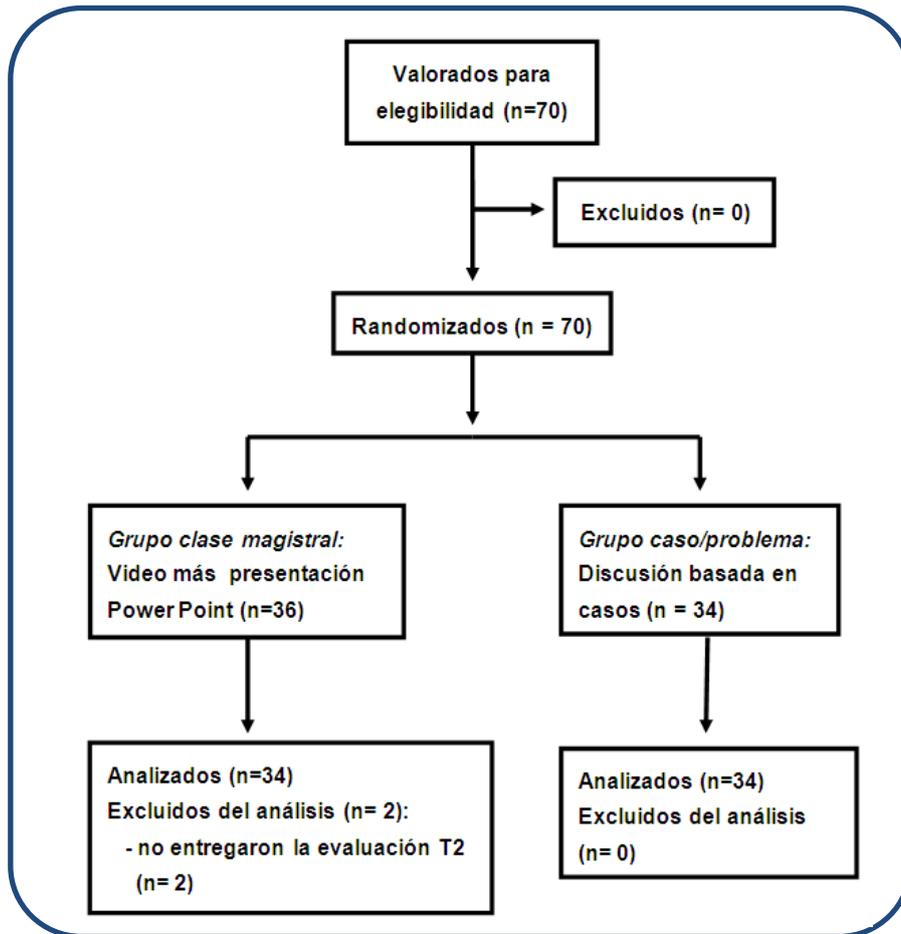


Figura 4. Diagrama de flujo que muestra la progresión de los estudiantes de medicina en el estudio.

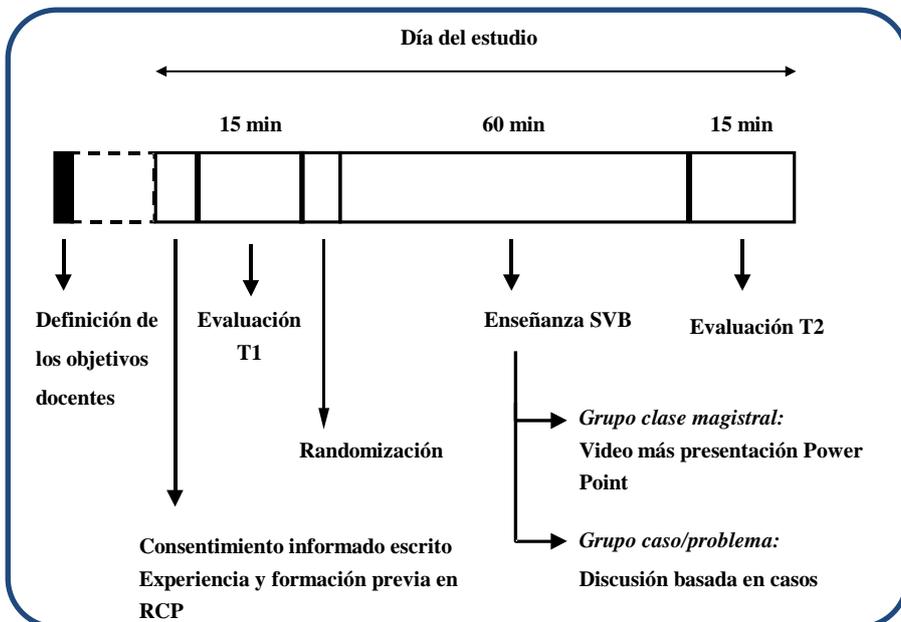
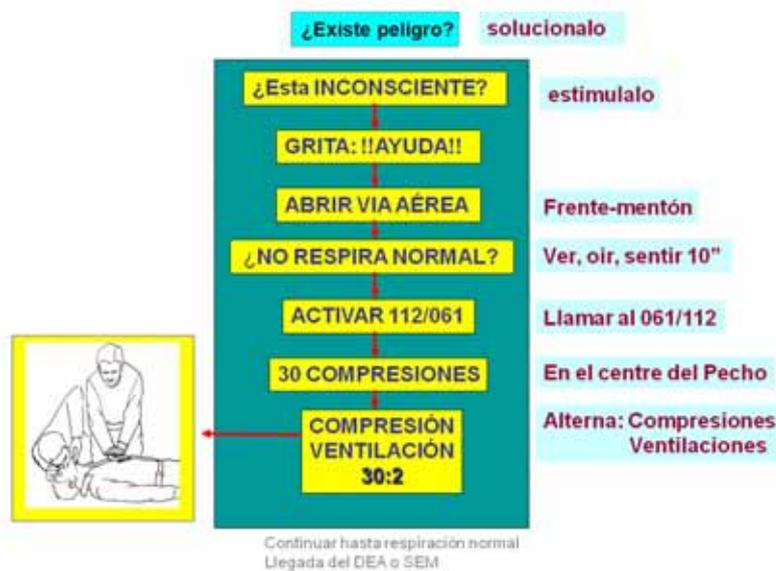


Figura 5. Esquema de la metodología seguida en el estudio de pregrado. Evaluación T1: test y grabación de video antes de la enseñanza de los algoritmos de soporte vital básico (SVB); evaluación T2: después de la enseñanza de los algoritmos de BLS; RCP: reanimación cardiopulmonar.

4.2.3. Enseñanza

La enseñanza se realizó en el Laboratorio de Habilidades Clínicas de la Facultad de Medicina de la Universidad de Barcelona. La estación de trabajo trató sobre la enseñanza del *soporte vital básico* de acuerdo a las Guías del *European Resuscitation Council* [133] (Figura 6). Los dos profesores que participaron en el estudio estaban acreditados como instructores por el *European Resuscitation Council* y tenían acumulada una experiencia de más de 5 años en la enseñanza del *soporte vital básico*.

A: sin desfibrilador semi-automático externo



B: con desfibrilador semi-automático externo.

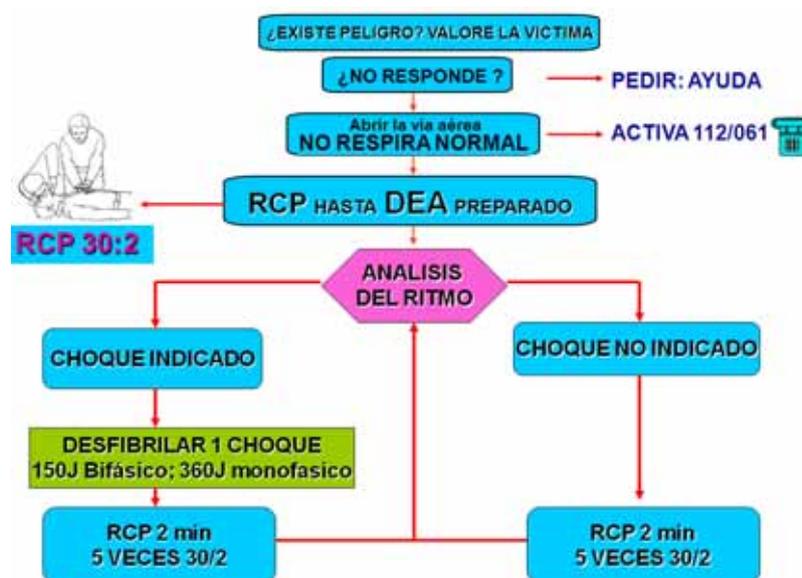


Figura 6. Secuencia de acción en soporte vital básico. Recomendaciones del *European Resuscitation Council* [133] RCP: reanimación cardiopulmonar, DEA: desfibrilador automático externo.

Los estudiantes del *grupo clase magistral* recibieron el contenido en un formato multimedia: 20 minutos de grabación de video doméstico sobre *soporte vital básico*, más 25 minutos de presentación Power Point y 15 minutos para preguntas y respuestas. La grabación de video y la presentación Power Point fueron complementarias. La grabación de video mostró escenas simuladas de *soporte vital básico* en el ámbito extrahospitalario ejecutadas por uno y dos reanimadores; se incluyeron todos los pasos de los algoritmos del *European Resuscitation Council*: detección de la víctima, seguridad de la víctima y de los reanimadores, valoración del nivel de conciencia, solicitud de ayuda, permeabilidad de la vía aérea, valoración de la respiración, telefonar al 112, compresiones torácicas, secuencia compresiones/ventilaciones y re-valoración de la víctima. La presentación Power Point incluyó la epidemiología, etiología y fisiopatología del paro cardíaco, las evidencias existentes sobre la eficacia de las maniobras de *soporte vital básico*, la vía aérea instrumentalizada, la ventilación con bolsa autoinflable-válvula-mascarilla, la posición lateral de seguridad y la actuación ante la obstrucción de la vía aérea por cuerpo extraño. Se organizaron cuatro estaciones de trabajo con 8-10 alumnos participando en cada una de ellas. Los recursos docentes fueron un ordenador para todos los estudiantes en cada estación de trabajo y la presencia constante del profesor.

Los estudiantes del *grupo caso/problema* empezaron directamente con la discusión de tres escenarios problema: 1) una víctima en paro cardíaco no traumático; 2) una víctima inconsciente con respiración espontánea efectiva y no efectiva; y 3) una víctima con obstrucción de la vía aérea por cuerpo extraño. Se organizaron cuatro estaciones de trabajo con 8-10 alumnos participando en cada una de ellas. El profesor ejecutaba las acciones sobre el maniquí durante la simulación de los casos escenario, interactuando con los alumnos. Los recursos docentes incluyeron un maniquí de *soporte vital básico* Resusci®Anne SkillReporter™ (Laerdal®), material de barrera para la ventilación boca-boca, cánulas orofaríngeas, y un sistema de bolsa reservorio-válvula-mascarilla. Para simular la posición lateral de seguridad y las maniobras de obstrucción de la vía aérea por cuerpo extraño, el profesor solicitó la colaboración voluntaria de un estudiante.

Después de la presentación multimedia en el *grupo clase magistral* y de la discusión de los casos en el *grupo caso/problema*, la estación se completó en ambos grupos con 90 minutos de prácticas de los alumnos con los maniqués utilizando el material de *soporte vital básico*. Los profesores corrigieron a los estudiantes durante estas prácticas. El objetivo del estudio no incluyó los efectos de las prácticas de los estudiantes sobre las habilidades cognitivas porque esta sesión docente podría haber producido un efecto techo, eliminando las diferencias potenciales entre los dos métodos de enseñanza.

4.2.4. Evaluaciones

Las evaluaciones incluyeron un test y una grabación de video doméstico (Tabla 6).

Tabla 6. Evaluaciones: preguntas, campos de habilidades cognitivas explorados, respuestas correctas y criterios de puntuación.

Test

Preguntas

1. "En la secuencia de soporte vital básico, que maniobras deberían realizarse ante una víctima adulta que yace en el suelo sobre su espalda, que no responde cuando se le estimula y que esta respirando con normalidad. Usted está solo y nadie acude después de gritar solicitando ayuda".
2. "En la secuencia de soporte vital básico, que maniobras deberían realizarse ante una víctima adulta que yace en el suelo sobre su espalda, que no responde cuando se le estimula. Usted está solo y nadie acude después de gritar solicitando ayuda".
3. "En la secuencia de soporte vital básico, que maniobras deberían realizarse ante una víctima adulta que yace en el suelo sobre su espalda, que no responde cuando se le estimula y que esta haciendo respiraciones agónicas ("gasping"). Usted está solo y nadie acude después de gritar solicitando ayuda. Usted ya ha telefonado al 112".

Campo medido:

1. "Decisión basada en la secuencia de soporte vital básico".

Respuestas correctas:

1. "Colocar a la víctima en la posición lateral de seguridad".
2. "Abrir la vía aérea".
3. "Compresiones torácicas".

Sistema de puntuación:

Un punto por cada respuesta correcta. Rango de puntuación: 0-3.

Grabación de video

Pregunta

"Usted está a punto de visualizar una grabación de video sobre la secuencia de soporte vital básico en adulto realizado por un solo reanimador. La secuencia de soporte vital básico que aparece en la escena puede contener o no maniobras realizadas incorrectamente. Por favor, indique las maniobras incorrectas, si las hubiera, que usted puede identificar".

Paso del soporte vital básico:

1. Seguridad.
2. Valorar inconsciencia.
3. Grita "ayuda".
4. Abrir la vía aérea.
5. Valorar si respira.
6. Activar 112.
7. Compresiones torácicas:
8. Relación compresiones : ventilaciones.
9. Parar para re-valorar a la víctima.

Video:

1. Víctima y reanimador seguros.
2. Estimular y gritar: no responde.
3. Ayuda!!
4. Maniobra frente-mentón.
5. Manteniendo la vía aérea abierta, mirar, escuchar y sentir: no respira.
6. Activar 112.
7. Talón de la mano izquierda localizado sobre la parte alta del abdomen de la víctima, dedos de ambas manos entrelazados, posición del reanimador vertical, brazos estirados, presión con descenso del tórax 4-5 cm, liberación de la presión sin perder el contacto entre las manos del reanimador y el tórax de la víctima. Frecuencia de compresiones 50 veces por minuto.
8. 5:2.
9. Después de 5 compresiones y 2 ventilaciones.

Campo medido:

1. "Detección de errores en la secuencia de soporte vital básico"

Respuestas correctas (errores detectados correctamente):

1. "Talón de la mano izquierda localizado sobre la parte alta del abdomen de la víctima". (guías del European Resuscitation Council: en el centro del tórax de la víctima).
2. "Frecuencia de las compresiones torácicas 50 veces por minuto". (guías del European Resuscitation Council: frecuencia aproximada de 100 min⁻¹).
3. "Relación compresiones : ventilaciones: 5:2". (guías del European Resuscitation Council: relación 30:2).
4. "Parar para re-evaluar a la víctima después de 5 compresiones y 2 ventilaciones". (guías del European Resuscitation Council: sólo si la víctima empieza a respirar con normalidad).

Sistema de puntuación:

Un punto por cada respuesta correcta. Rango de puntuación: 0-4.

El test incluyó tres preguntas de respuesta abierta sobre tres escenarios clínicos. Los estudiantes tenían que responder que maniobra de *soporte vital básico* estaba indicada en cada caso. La escena de video mostraba cuatro acciones erróneas en la ejecución de la secuencia de *soporte vital básico* ante un paro cardíaco ejecutadas por un solo reanimador sobre un maniquí adulto (Resusci[®] Anne). Los estudiantes tenían que identificar las acciones incorrectas y escribirlas en una plantilla. Las evaluaciones no analizaron las habilidades de ejecución de los estudiantes.

Seis meses antes del estudio se realizó una *prueba piloto*. Veintinueve estudiantes de dos estaciones de trabajo en *soporte vital básico* completaron las evaluaciones, el test y el video, y fueron puntuados por los dos mismos evaluadores que posteriormente puntuarían a los participantes del estudio. Los evaluadores confirmaron que las evaluaciones eran comprensibles y podían completarse en 15 minutos. Este estudio piloto sirvió para excluir los datos del test y el video que causaban discrepancia entre los evaluadores y para preparar el diseño final del test y del video.

Las evaluaciones se cumplimentaron antes (*evaluación pre = evaluación T1*) e inmediatamente después (*evaluación post = evaluación T2*) de la aplicación de los dos métodos de enseñanza de los algoritmos de *soporte vital básico* (Figura 5). La *evaluación post = evaluación T2* se cumplimentó antes de los 90 minutos de prácticas de los estudiantes sobre maniqués para evitar la posible influencia de ésta última sobre los resultados. Los estudiantes tuvieron un máximo de 15 minutos para completar las evaluaciones y no se les permitió realizar comentarios sobre las evaluaciones hasta que finalizó la *evaluación post = evaluación T2*.

Los objetivos primarios valorados fueron dos campos diferentes de conocimiento en *soporte vital básico*: campo 1 “*decisiones basadas en la secuencia de soporte vital básico*” (test) y campo 2 “*detección de errores en la ejecución del soporte vital básico*” (grabación de video). Contamos el número total de respuestas correctas en las *evaluaciones pre = evaluación T1* y *post = evaluación T2*. Cada respuesta correcta (*ítem*) puntuó 1 punto y las respuestas incorrectas o preguntas no contestadas puntuaron 0 puntos.

Los objetivos secundarios fueron el incremento absoluto de puntuación entre los diferentes tiempos de cada evaluación (valor final – valor inicial) y la ganancia de puntuación (Δ), definida como el cociente entre el incremento absoluto en la puntuación y el incremento potencial:

$$\Delta \text{ Test} = 100 * [(\text{valor final} - \text{valor inicial}) / (3 - \text{valor inicial})]$$

$$\Delta \text{ Video} = 100 * [(\text{valor final} - \text{valor inicial}) / (4 - \text{valor inicial})]$$

Por último, contamos el número de estudiantes que mejoraron sus puntuaciones después de la enseñanza (T2) respecto a sus puntuaciones antes de la enseñanza (T1); consideramos relevante un porcentaje > 50%.

4.3. Metodología Artículo 2. Carrero E, Gomar C, Penzo W, Rull M. Comparison between lecture-based approach and case/problem-based learning discussion for teaching pre-anaesthetic assessment. *Eur J Anaesthesiol* 2007; 24: 1008-15.

4.3.1. Diseño

El diseño del estudio en *postgrado* se adaptó al programa de formación teórica de los residentes de Anestesiología de primer año de la Societat Catalana d'Anestesiologia, Reanimació i Terapèutica del Dolor (SCARTD). La clase "valoración pre-anestésica" tiene asignada una hora de duración y se expone tradicionalmente como *clase magistral* utilizando una presentación Power Point de 50 minutos de duración más 10 minutos finales para preguntas de los residentes. La evaluación de los residentes es teórica y está incluida como preguntas tipo test de respuesta múltiple en el examen final del curso académico.

El diseño de esta parte del estudio cumplió las características de la metodología general del estudio (Apartado 4.1.1.).

4.3.2. Sujetos

4.3.2.1. Cálculo del tamaño de la muestra

El tamaño de la muestra en esta parte del estudio estuvo determinado por el número de residentes de Anestesiología de primer año asistentes a la clase "valoración pre-anestésica" de la SCARTD, que autorizaron su participación en el estudio y que cumplimentaron las *evaluaciones pre y post*. El estudio se inició el curso académico 1999/2000 coincidieron con el periodo de incorporación de los residentes de primer año a las clases de la SCARTD. En 1999 había 37 residentes de primer año en Cataluña, tamaño que consideramos insuficiente para validar nuestro estudio. Tampoco teníamos datos previos sobre el porcentaje de asistencia a clase; estimamos que éste sería parecido al porcentaje de asistencia de los residentes de segundo año que oscilaba entorno al 75%. En este contexto la única forma de incrementar el tamaño de la muestra era repetir el estudio el siguiente año. Los autores consideramos que un máximo de dos años académicos consecutivos no incorporaban cambios curriculares en la formación previa de los residentes ni modificaciones en la metodología y experiencia del profesorado que pudieran sesgar los resultados.

4.3.2.2. Población del estudio

La población del estudio incluyó a los residentes de Anestesiología de primer año de Cataluña de dos años consecutivos, 1999 y 2000. La Figura 7 muestra el diagrama de randomización y participación. El estudio tuvo lugar durante 2 días de 2 años consecutivos durante la clase "valoración pre-anestésica" del programa de formación teórica para residentes de Anestesiología de primer año de la SCARTD. La Figura 8 muestra una representación esquemática del estudio.

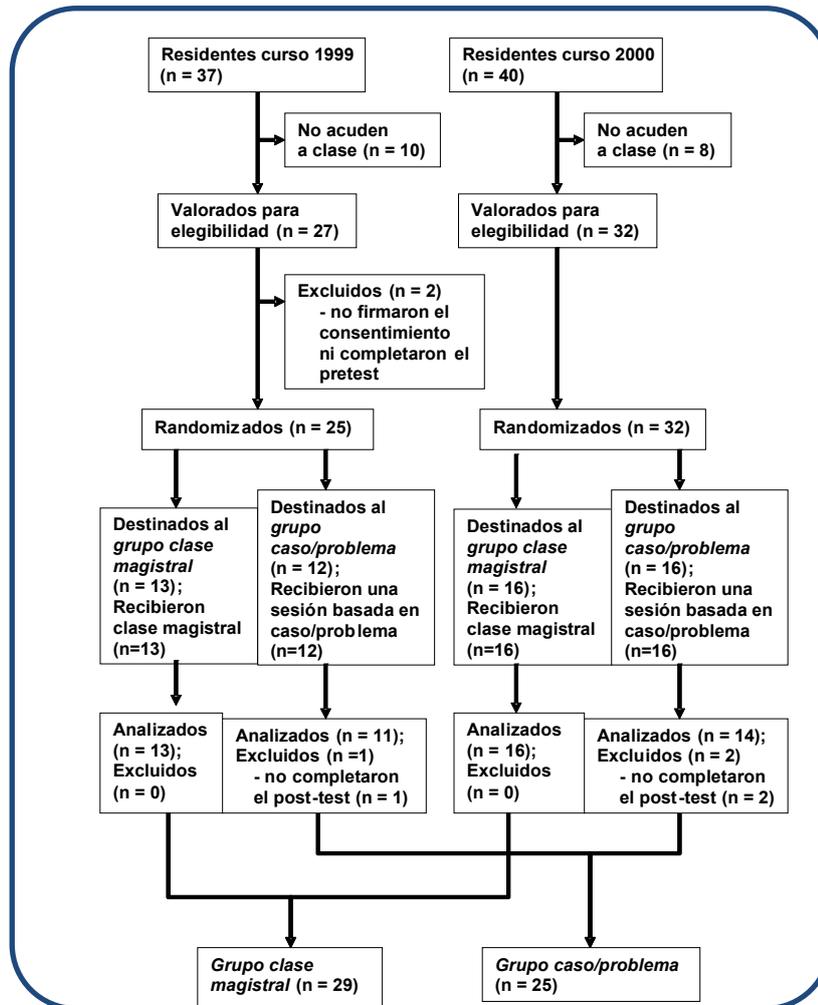


Figura 7. Diagrama de flujo que muestra la progresión de los residentes de Anestesiología en el estudio. Pre-test: evaluación antes de la sesión docente; post-test: evaluación después de la sesión docente.

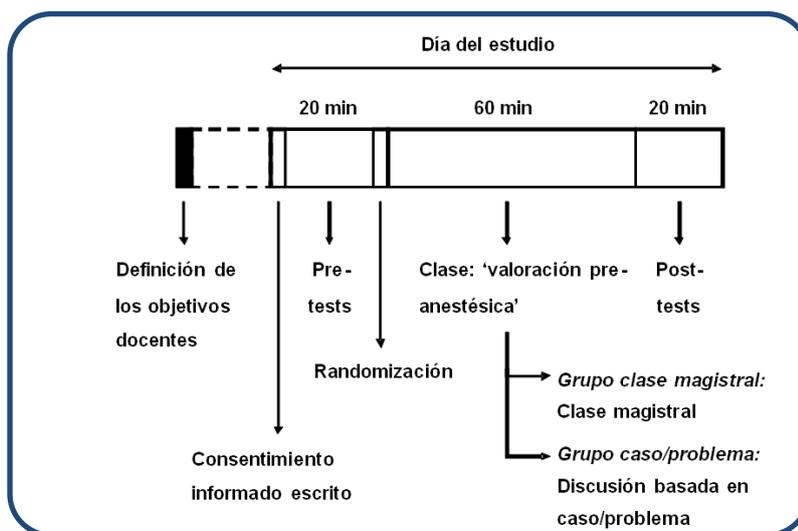


Figura 8. Esquema de la metodología seguida en el estudio de postgrado. Pre-test: evaluación antes de la sesión docente; post-test: evaluación después de la sesión docente.

4.3.3. Enseñanza

La enseñanza se realizó en las aulas de la Academia de Ciencias Médicas de la SCARTD. La clase trató sobre la *valoración pre-anestésica* de los pacientes que van a ser sometidos a una intervención quirúrgica. La Tabla 7 resume los contenidos principales de la clase.

Tabla 7. Contenidos docentes de la clase valoración pre-anestésica.

1. Evaluación del riesgo perioperatorio y medidas para reducir la morbilidad y la mortalidad.
2. Análisis de la información obtenida en la historia clínica.
3. Exploración física y estudios complementarios.
4. Valoración de las dificultades para el manejo de la vía aérea.
5. Clasificación Mallampati [146] de la orofaringe para la predicción de dificultad de intubación traqueal
6. Clasificación ASA [147] del estado físico.
7. Optimización del paciente antes de la cirugía
8. Estrategia anestésica
9. Información al paciente y consentimiento informado
10. Premedicación
11. Analgesia y cuidados postoperatorios

ASA: American Society of Anesthesiologists

La sesión docente duró 60 minutos en ambos grupos. La exposición en el *grupo clase magistral* fue de 50 minutos más 10 minutos para preguntas, mientras el *grupo caso/problema* tuvo 60 minutos para la discusión del caso. La *clase magistral* consistió en una exposición estructurada según los objetivos docentes a partir de una presentación Power Point. Su formato incorporó esquemas y ejemplos aclaradores; los residentes no participaron durante la clase excepto en los 10 últimos minutos de preguntas. El caso problema discutido en el *grupo caso/problema* fue un paciente con una historia clínica de elevado riesgo preoperatorio, con medicación concomitante que precisaba ser reajustada para la cirugía y vía aérea difícil; el paciente estaba programado primero para cirugía urológica electiva, y meses después, precisó cirugía urgente por shock traumático.

4.3.4. Evaluaciones

Se diseñaron cuatro tests, basados en cuatro casos clínicos similares que, a pares, cumplían los objetivos docentes (Tabla 8). Los casos clínicos simulaban cuatro pacientes programados para cirugía. Los tests fueron validados 4 meses antes del estudio por dos anesthesiólogos de plantilla, tres residentes de Anestesiología de segundo año, un residente de Anestesiología de tercer año y otro de cuarto año, y fueron puntuados por los dos mismos evaluadores quienes posteriormente puntuarían a los participantes del estudio. Los evaluadores confirmaron que los tests eran comprensibles, tenían el mismo grado de dificultad y podían ser cumplimentados en 20 minutos. Esta validación se hizo para excluir los datos o preguntas en los tests que causaran desacuerdo entre los evaluadores, y así generar los cuatro tests finales.

Tabla 8. Evaluaciones antes de la sesión docente (pre-tests) y después de la sesión docente (post-tests): preguntas, campos de conocimiento medidos, respuestas correctas y criterios de puntuación.

Pre-tests

Test 1: Paciente de 84 años con antecedentes patológicos de hipertensión arterial no controlada e insuficiencia cardiaca congestiva. Hace un mes se le practicó una laparotomía exploradora por un cuadro de abdomen agudo sufriendo un infarto de miocardio en el postoperatorio inmediato. Presenta evisceración infraumbilical que requiere intervención quirúrgica urgente.

Test 2: Mujer de 60 años, 155 cm, 60 kg. Alérgica al látex. Antecedentes patológicos de asma bronquial en tratamiento con broncodilatadores inhalados y prednisona 10 mg al día desde hace años. Exploración física: apertura bucal limitada, solo se observa lengua y paladar duro. Gasometría basal: PaO₂ 56 mm Hg. Se programa para cirugía de hernia de hiato.

Post-tests

Test 3: Embarazada a término, se requiere cesarea por retraso en el trabajo de parto. Antecedentes familiares de neoplasia de mama. Fumadora de 1 paquete al día los últimos 10 años. Apendicectomizada hace 4 años. La exploración de Mallampati muestra paladar duro y una pequeña parte de paladar blando sin que se vean los pilares ni la úvula.

Test 4: Paciente de 50 años sin hábitos tóxicos. Sustitución valvular aórtica hace dos años, con buen resultado funcional; desde entonces en tratamiento con anticoagulantes orales. Se requiere cirugía urgente por fractura abierta de tibia y peroné. La exploración física revela palidez cutáneo-mucosa, taquicardia 130 latidos por minuto y presión arterial 75 / 50 mmHg. Mallampati clase I.

Pregunta 1

"Subraye los ítems en el texto que, en su opinión, tienen implicaciones para la planificación de la anestesia".

Campo medido: Campo 1, "Reconocimiento de datos clínicos con implicaciones anestésicas".

Respuestas correctas (test 1): "edad 84", "hipertensión", "insuficiencia cardiaca", "infarto de miocardio", "evisceración infraumbilical".

Respuestas correctas (test 2): "Alergia al latex", "asma", "prednisona", "apertura bucal limitada", "PaO₂ 56 mmHg", "hernia de hiato".

Respuestas correctas (test 3): "Embarazada a término", "cesarea", "fumadora", "exploración de Mallampati".

Respuestas correctas (test 4): "sustitución valvular aortica", "anticoagulantes orales", "fractura abierta de tibia y peroné", "palidez", "taquicardia", "presión arterial 75/50 mmHg", "Mallampati clase I".

Sistema de puntuación para pre y post-tests: 1 punto por cada ítem subrayado correctamente. Rango de puntuación: 0-11.

Pregunta 2

"Por cada ítem subrayado, explique porque piensa usted que es relevante para la anestesia".

Campo medido: Campo 2, "Razonamiento de datos clínicos con implicaciones anestésicas"

Respuestas correctas: un razonamiento correcto por cada uno de los ítems subrayados correctamente.

Sistema de puntuación para pre y post-tests: 1 punto por cada ítem subrayado y razonado correctamente. Rango de puntuación: 0-11.

Pregunta 3

"Evalúe el riesgo del paciente de acuerdo a la clasificación ASA".

Campo medido: Campo 3, "Clase ASA".

Respuesta correcta (test 1): 'ASA clase IV'.

Respuesta correcta (test 2): 'ASA clase III'.

Respuesta correcta (test 3): 'ASA clase II'.

Respuesta correcta (test 4): 'ASA clase III'.

Sistema de puntuación para pre y post-tests: 1 punto por cada ítem contestado correctamente. Rango de puntuación: 0-2.

Pregunta 4 (solo para tests 2 y 3):

"¿Cuál es la clasificación Mallampati del paciente?"

Campo medido: Campo 4, "Clase Mallampati".

Respuesta correcta (test 2): 'Mallampati clase IV'.

Respuesta correcta (test 3): 'Mallampati clase III'.

Sistema de puntuación para pre y post-tests: 1 punto por cada ítem contestado correctamente. Rango de puntuación: 0-1.

Pregunta 5 (solo para test 1 y test 3):

"¿Qué técnica anestésica (general o locorregional) elegiría en este caso?"

Campo medido: Campo 5 "Elección de la técnica anestésica"

Respuesta correcta (test 1): "Anestesia loco-regional"

Respuesta correcta (test 3): "Anestesia loco-regional"

Sistema de puntuación para pre y post-tests: 1 punto por cada ítem contestado correctamente. Rango de puntuación: 0-1.

Pregunta 6 (para test 1 y test 3):

"Justifique su elección de la técnica anestésica".

Pregunta 6 (para test 2 y test 4):

"¿Por qué cree que la mayor técnica anestésica para este paciente es la anestesia general?"

Campo medido: Campo 6, "Razonamiento de la elección de la técnica anestésica"

Respuestas correctas: Un razonamiento correcto para indicar anestesia regional (para test 1 y test 3) o anestesia general (para test 2 y test 4).

Sistema de puntuación para pre y post-tests: 1 punto por cada ítem razonado correctamente. Rango de puntuación: 0-2.

El primer día, se decidió aleatoriamente, con los residentes del primer año del estudio, que par de tests cumplimentaría cada grupo antes de la sesión docente (tests 1 y 2 = *pre-tests*) y cuales después de la sesión docente (tests 3 y 4 = *post-tests*). El orden se repitió para los residentes del segundo año del estudio. El tiempo máximo para cumplimentar los *pre-tests* y *post-tests* fue 20 minutos. Los *pre-tests* y *post-tests* se realizaron inmediatamente antes y después de la sesión docente (Figura 8). No se permitió a los residentes realizar ningún comentario relativo a los tests hasta que se finalizaron los *post-tests*.

Las variables medidas en los tests fueron seis campos de conocimiento en el tema “*valoración pre-anestésica*”: campo 1, “*reconocimiento de los datos clínicos con implicaciones anestésicas*”; campo 2, “*razonamiento de los datos clínicos con implicaciones anestésicas*”; campo 3, “*clase ASA: clasificación del estado físico de la American Society of Anesthesiologists (ASA)*”; campo 4, “*Clase Mallampati: clasificación Mallampati de la orofaringe para predecir la dificultad de intubación orotraqueal*”; campo 5, “*elección de la técnica anestésica*”; campo 6, “*razonamiento de la elección de la técnica anestésica*”.

Cada campo tenía un número fijo de *ítems* sobre los que se contabilizó el número total de respuestas correctas (Tabla 8). Cada *ítem* respondido correctamente puntuó 1 punto y las respuestas incorrectas o las preguntas sin contestar puntuaron 0 puntos. Calculamos la puntuación obtenida para cada uno de los campos en los *pre-tests* y *post-tests* y la diferencia de puntuaciones entre los *post-tests* y *pre-tests*. Establecimos entonces si existían diferencias en las puntuaciones entre los grupos e intragrupo, y si las diferencias en los resultados podrían haber sido causadas por el año de estudio considerado.

4.4. Metodología Artículo 3. Carrero E, Gomar C, Fábregas N, Penzo W, Castillo J, Villalonga A. Clase magistral versus aprendizaje basado en caso/problema para la enseñanza del embolismo aéreo en formación médica continuada. *Rev Esp Anestesiol Reanim* 2008; 55: 202-9.

4.4.1. Diseño

El diseño del estudio en *Formación Médica Continuada* se adaptó al programa del Curso 5 del Comité para la Educación Europea en Anestesiología (CEEAA), Centro de Cataluña. El programa del curso trata temas relacionados con el sistema nervioso, la anestesia regional y el tratamiento del dolor. La clase “*embolismo aéreo*” está incluida dentro de la anestesia en neurocirugía [140,141]. Su duración es de 40 minutos. Tradicionalmente se expone como *clase magistral* en forma de presentación Power Point de 30 minutos de duración más 10 minutos para preguntas y respuestas. La evaluación de los participantes es teórica y está incluida como preguntas tipo test de respuesta múltiple en la autoevaluación anónima que se hace al final del curso.

El diseño de esta parte del estudio cumplió las características de la metodología general del estudio (Apartado 4.1.1.).

4.4.2. Sujetos

4.4.2.1. Cálculo del tamaño de la muestra

Para el cálculo del tamaño de la muestra se estimó una eficacia media del 75% ($p_1=0,75$); se consideró relevante una diferencia en las puntuaciones del 25% ($p_2=1$) y se estimó una pérdida de casos del 10% ($R=0,1$). El tamaño de la muestra resultante fue de 29 participantes en cada grupo.

4.4.2.2. Población del estudio

La población del estudio fueron 59 anesthesiólogos en ejercicio profesional inscritos en el Curso 5 del Centro CEEAA de Cataluña. La Figura 9 muestra el diagrama de randomización y participación de los alumnos en el estudio y la Figura 10 las etapas del estudio.

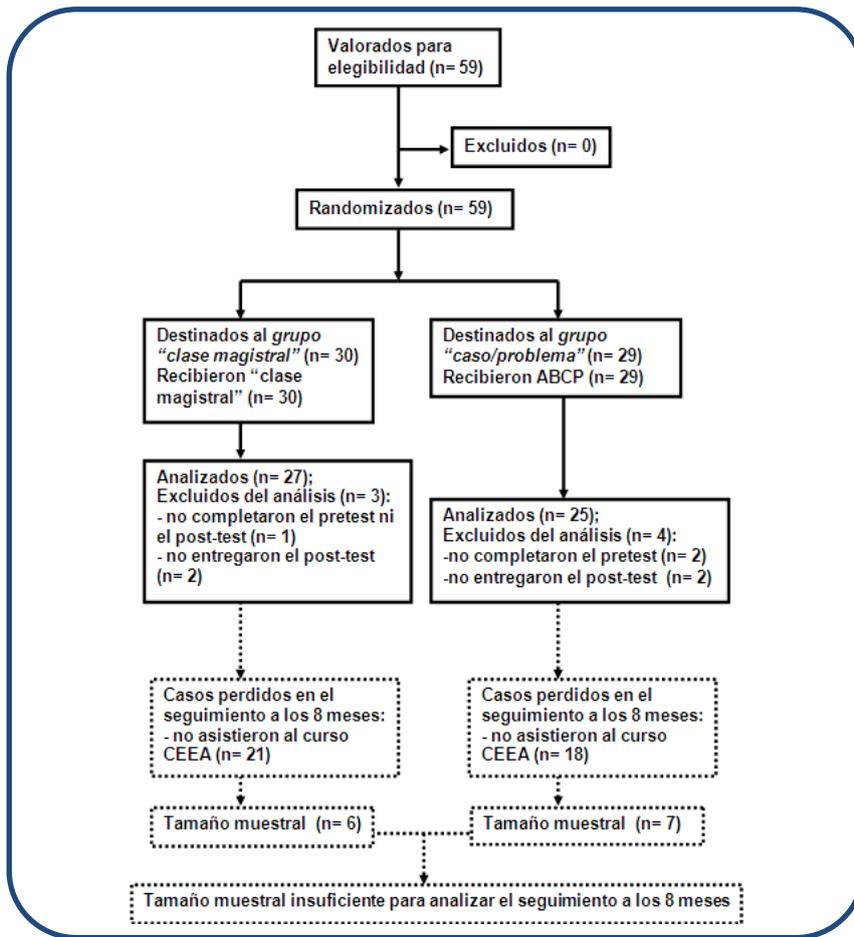


Figura 9. Diagrama de flujo que muestra la progresión de los participantes del curso del Comité para la Educación Europea en Anestesiología (CEEa) en el estudio. ABCP: aprendizaje basado en caso/problema. Pretest: evaluación antes de la enseñanza; post-test: evaluación después de la enseñanza.

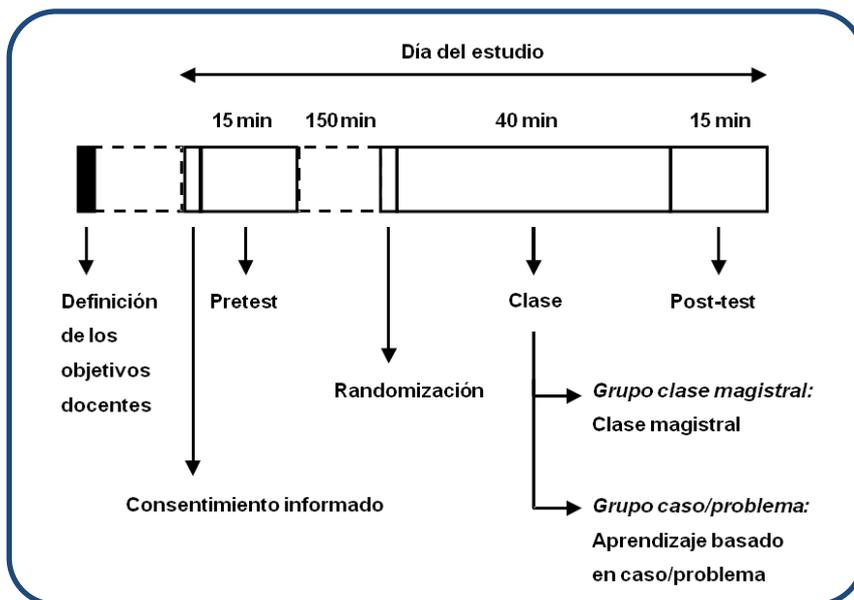


Figura 10. Esquema de la metodología seguida en el estudio de postgrado. Pre-test: evaluación antes de la enseñanza; post-test: evaluación después de la enseñanza.

4.4.3. Enseñanza

La enseñanza se realizó en las aulas que el Centro *CEEA* de Cataluña tiene habilitadas para los cursos de *formación continuada*. El contenido de la enseñanza fue el *embolismo aéreo* en anestesia. Los objetivos docentes (Tabla 9) se definieron de acuerdo a la literatura especializada [139-141,143].

Tabla 9. Objetivos docentes de la clase embolismo aéreo en anestesia.

1. Definición de embolismo aéreo
2. Incidencia
3. Factores de riesgo
4. Fisiopatología
5. Manifestaciones clínicas
6. Diagnóstico
7. Monitorización
8. Pronóstico
9. Prevención
10. Tratamiento
11. Embolismo areo paradójico

Los dos profesores que impartieron la clase eran miembros del Centro *CEEA* de Cataluña. Ambas clases duraron 40 minutos. En el *grupo clase magistral* la clase incluyó 30 minutos de *lección magistral* más 10 minutos para preguntas y respuestas, en el *grupo caso/problema* los 40 minutos fueron de discusión. El material de soporte docente fue una presentación Power Point en el *grupo clase magistral* y una pizarra en el *grupo caso/problema*. La *clase magistral* se estructuró conforme a los objetivos docentes e incorporó fotografías y esquemas para facilitar la exposición. El caso clínico discutido en el *grupo caso/problema* fue un paciente en respiración espontánea que presentó un episodio de *embolismo aéreo* durante la cirugía de su enfermedad de Parkinson [148].

4.4.4. Evaluaciones

La evaluación se basó en las respuestas a dos tests referidos a dos casos clínicos hipotéticos (Tabla 10) que cumplieran los mismos objetivos docentes. Ambos casos simulaban un *embolismo aéreo* durante un procedimiento neuroquirúrgico. No se ha descrito un sistema de evaluación y puntuación estandarizado para el *embolismo aéreo*. El formato y contenido de los tests fueron similares al sub-algoritmo del *embolismo aéreo* validado por Williamson et al. [139] para el manejo de situaciones críticas en anestesia. Previamente al estudio, se realizó una *prueba piloto* en un grupo de anesestesiólogos que no participaron como alumnos en el curso; los tests fueron puntuados por los dos mismos evaluadores que posteriormente puntuaron a los participantes del estudio. Se confirmó que los tests eran comprensibles, del mismo grado de dificultad y fáciles de completar en 15 minutos. Esta prueba piloto sirvió para eliminar de los tests los *items* o preguntas que causaban discrepancia entre los profesores. El mismo día del estudio se decidió al azar cual de los dos tests se tenía que cumplimentar antes de la enseñanza (*pretest*) y cual después de la enseñanza (*post-test*).

Tabla 10. Evaluación antes (pre-test) y después de la enseñanza (post-test).**Pretest**

Mujer de 49 años, peso: 60 Kg, altura: 159 cm; ingesta de alcohol: 20 g día⁻¹; hipertensa en tratamiento con captopril. Diagnosticada por escáner y resonancia magnética cerebrales de meningioma del ángulo ponto-bulbo-cerebeloso; inició tratamiento con corticoides y diuréticos. Se programa para cirugía de la fosa posterior en posición sentada. Mantenimiento anestésico: propofol, fentanilo y vecuronio. Craneotomía sin complicaciones. Durante la exéresis del tumor, coincidiendo con un episodio de sangrado venoso y de forma brusca, se objetivó un descenso en el CO₂ espirado (ETCO₂) a 6 mmHg, taquicardia e hipotensión arterial.

1. "Por favor, marque los ítems que, en su opinión, tienen relación con la complicación expuesta y justifique por qué".
2. "Establezca el diagnóstico diferencial. Señale su diagnóstico de sospecha".
3. "A su juicio, qué monitorización sería útil para confirmar su diagnóstico de sospecha".
4. "Escriba que medidas terapéuticas adoptaría ante la complicación descrita en el caso clínico".

Post-test

Varón de 55 años, peso: 80 Kg, altura 177 cm; fumador de un paquete de cigarrillos al día desde hace 20 años. Intolerancia gástrica a la aspirina. Colectomía laparoscópica hace 2 años. Se programa para fijación cervical instrumentalizada por inestabilidad cervical traumática; abordaje posterior con cabeza incorporada. Se cateterizan: vena subclavia derecha, 2 venas periféricas y arteria radial izquierda. El paciente permanece estable las 2 primeras horas de la intervención. Sangrado en 2 horas: 300 mL; líquidos perfundidos en 2 horas: 1L de cristaloides y 0,5 L de coloides. A las 2 horas y 5 minutos del inicio de la intervención se registró de forma súbita: marcada depresión del segmento ST; descenso del ETCO₂ a 10 mmHg en un minuto hasta llegar finalmente a 0 mmHg; descenso de la presión arterial a valores indetectables y disminución de las cifras de SpO₂ hasta perderse el registro.

1. "Por favor, marque los ítems que, en su opinión, tienen relación con la complicación expuesta y justifique por qué".
2. "Establezca el diagnóstico diferencial. Señale su diagnóstico de sospecha".
3. "A su juicio, qué monitorización sería útil para confirmar su diagnóstico de sospecha".
4. "Escriba que medidas terapéuticas adoptaría ante la complicación descrita en el caso clínico".

Antes de la clase, no se proporcionó a los participantes ningún material docente del tema "embolismo aéreo", a diferencia del resto de temas del curso. El *pretest* se cumplimentó 150 minutos antes de la clase y el *post-test* se realizó inmediatamente después de la clase. El tiempo máximo permitido para la cumplimentación de cada test fue de 15 minutos. Los participantes solo pudieron realizar preguntas relacionadas con los casos clínicos de los tests cuando finalizó el *post-test*.

Las variables medidas en los tests fueron cuatro campos de conocimiento sobre el tema "embolismo aéreo en anestesia": 1) *Factores de riesgo y síntomas*; 2) *Diagnóstico*; 3) *Monitorización*; y 4) *Tratamiento*. Cada campo incluía un número determinado de *ítems* sobre los que se contabilizó el número total de respuestas correctas (Tabla 11). Los *ítems* contestados correctamente puntuaron 1 punto y los *ítems* no contestados o contestados incorrectamente puntuaron 0 puntos. Se calculó la puntuación obtenida para cada uno de los campos en el *pretest* y en el *post-test*. También se calculó la diferencia en la puntuación "*post-test* menos *pretest*". Por último, se analizó si existían diferencias entre los grupos e intragrupo.

Para evaluar la retención de conocimientos, se intentó repetir el *pretest* 8 meses más tarde, coincidiendo con el inicio de otro curso en el que se preveía la inscripción de gran parte de los participantes. Sin embargo, sólo 14 participantes se inscribieron en el nuevo curso (6 pertenecientes

al grupo clase magistral y 8 pertenecientes al grupo caso/problema) por lo que los datos resultaron insuficientes para su análisis.

Tabla 11. Preguntas, respuestas correctas, criterios y rangos de puntuación correspondientes a los cuatro campos de conocimiento medidos en los tests.

| | |
|---|---|
| Campo 1: Factores de riesgo y síntomas. "Por favor marque los ítems que, en su opinión, tienen relación con la complicación sujeta y justifique por qué" | |
| Respuestas correctas | Pretest "Posición ventral"; "colocándose con un episodio de segundo ventos"; "de forma lateral"; "descansar en el ETCD ₂ a 6 mmHg"; "taquicardia"; "hipotensión". Post-test "Abordaje postural con cabeza inclinada"; "de forma lateral"; "mancha de posición del segmento ST"; "descenso del ETCD ₂ a 10 mmHg en un minuto hasta llegar finalmente a 0 mmHg"; "descenso de la presión arterial a valores indetectables"; "alteración de las ondas de SpO ₂ hasta perderse el registro" |
| Puntuación | 1 punto por cada ítem correcto más 1 punto por cada ítem correcto con comentario correcto |
| Rango de puntos | 0-12 |
| Campo 2: Diagnóstico. "Escriba el diagnóstico diferencial. Señale su diagnóstico de sospecha". | |
| Respuestas correctas | Pretest "Embolicismo aéreo" Post-test "Embolicismo aéreo" |
| Puntuación | 1 punto por incluir "embolicismo aéreo" en el diagnóstico diferencial más 1 punto por señalar "embolicismo aéreo" como diagnóstico de sospecha. |
| Rango de puntos | 0-2 |
| Campo 3: Monitorización. "A su juicio, qué monitorización sería útil para confirmar su diagnóstico de sospecha". | |
| Respuestas correctas | Pretest "Cinegrafía"; "ecocardiografía"; "fluoroscopia o doppler paravertebral"; "cablear aurícula derecha" Post-test "Cinegrafía"; "ecocardiografía"; "fluoroscopia o doppler paravertebral"; "cablear aurícula derecha" |
| Puntuación | 1 punto por cada ítem correcto |
| Rango de puntos | 0-4 |
| Campo 4: Tratamiento. "Escriba qué medidas terapéuticas adoptaría ante la complicación descrita en el caso clínico". | |
| Respuestas correctas | Pretest "Artes, lugar: entrada de aire, horizontal"; "para N ₂ O, incrementar FIO ₂ "; "soporte hemodinámico"; "expirar aire por catéter venoso central"; "colocar posición del paciente" Post-test "Artes, lugar: entrada de aire, horizontal"; "para N ₂ O, incrementar FIO ₂ "; "soporte hemodinámico"; "expirar aire por catéter venoso central"; "colocar posición del paciente" |
| Puntuación | 1 punto por cada ítem correcto |
| Rango de puntos | 0-5 |

5. RESULTADOS

5.1. Resultados Artículo 1. Carrero E, Gomar C, Penzo W, Fabregas N, Valero R, Sanchez-Etayo G. Teaching basic life support algorithms by either multimedia presentations or case based discussion equally improves the level of cognitive skills of undergraduate medical students. *Med Teach* 2009; Feb 24: e1-e7.

Disponible en: URL; <http://dx.doi.org/10.1080/01421590802512896>

Sesenta y ocho estudiantes de medicina completaron el estudio, 34 en cada grupo. Dos participantes en el *grupo clase magistral* fueron excluidos porque no completaron la evaluación después de la enseñanza (Figura 4). El porcentaje de participación en el estudio fue 97,14%.

La formación en *soporte vital básico* previa a la estación de trabajo fue similar en ambos grupos (Tabla 12, Figura 11).

Tabla 12. Características personales y experiencia formativa en reanimación cardiopulmonar (RCP) de los estudiantes previa al estudio.

| Datos demográficos | Grupo clase magistral (n = 34) | Grupo caso/problema (n = 34) | P |
|---|-----------------------------------|---------------------------------|------|
| Mujeres / hombres ^(a) | 27/7 (79.4/20.6) | 26/8 (76.5/23.5) | 0.77 |
| Edad ^(b) | 20 (20-21) | 20 (20-21) | 0.65 |
| Estudiantes que habían recibido formación previa en RCP | | | |
| - clases más prácticas ^(a) | 8 (23.5) | 11 (32.4) | 0.42 |
| - prácticas ^(a) | 1 (2.9) | 2 (5.9) | 1 |
| Estudiantes que habían realizado RCP en pacientes o maniqués ^(a) | 10 (29.4) | 13 (38.2) | 0.44 |
| Número de RCP realizadas ^(b) | 0 (0-0.75) | 0 (0-1) | 0.43 |

Datos expresados como ^(a) número de estudiantes (% en cada grupo) o ^(b) mediana (percentiles 25-75).

Test de Chi cuadrado (variables cualitativas), Test T (edad); Test U de Mann Whitney (número de RCP realizadas).

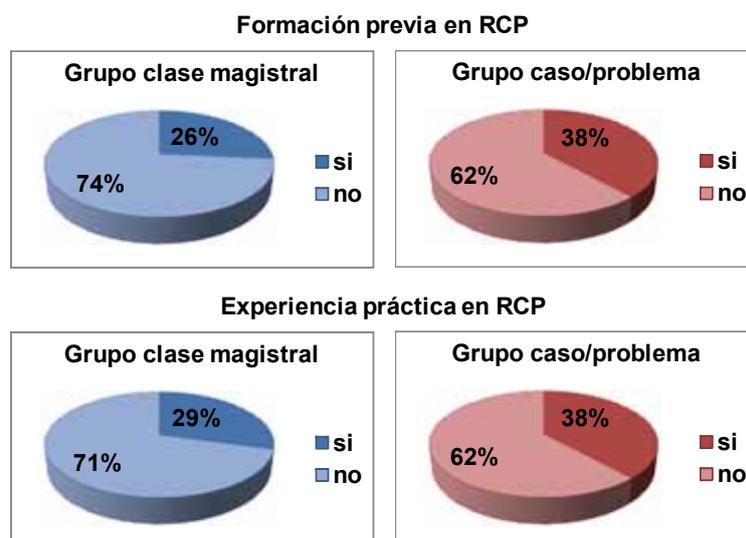
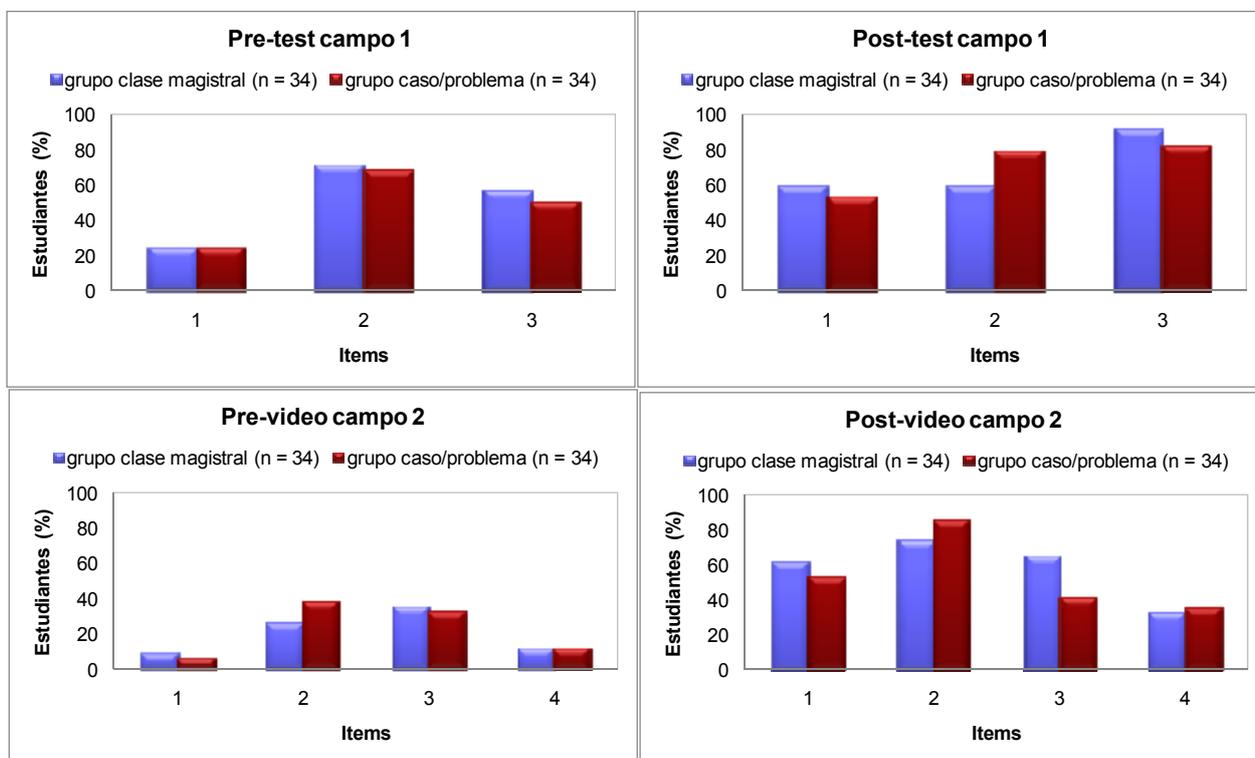


Figura 11. Formación y experiencia práctica previa en reanimación cardiopulmonar (RCP) de los estudiantes en cada uno de los grupos del estudio.

El porcentaje de alumnos que contestaron correctamente los 3 *ítems* del campo 1 “Decisiones basadas en la secuencia de soporte vital básico” y los 4 *ítems* del campo 2 “Detección de errores en la secuencia de soporte vital básico” fueron iguales antes y después de la enseñanza (Figura 12). No encontramos diferencias significativas entre los grupos, para ninguno de los *ítems* del campo 1 y campo 2, ni antes de la enseñanza ni después de la enseñanza.



Pre-test: evaluación (test) antes de la enseñanza; Post-test: evaluación (test) después de la enseñanza. Pre-video: evaluación (video) antes de la enseñanza; Post-video: evaluación (video) después de la enseñanza. Campo 1 “Decisiones basadas en la secuencia de soporte vital básico”; campo 2 “Detección de errores en la ejecución de la secuencia de soporte vital básico”.

Figura 12. Porcentaje de estudiantes que acertaron los ítems de cada uno de los campos de conocimiento medidos antes y después de la enseñanza de los algoritmos de soporte vital básico.

5.1.4. Puntuación obtenida en los campos de conocimiento medidos.

En la tabla 13 están resumidos los resultados de las evaluaciones. El índice *kappa* de concordancia entre evaluadores fue > 0.8 para todas las evaluaciones. No encontramos diferencias significativas entre los grupos en ninguna de las evaluaciones. En ambos grupos las puntuaciones después de la enseñanza mejoraron significativamente tanto en el test como en la grabación de video (valores $P < 0,001$). La mayoría de los estudiantes no obtuvieron la máxima puntuación para ninguna de las evaluaciones después de la enseñanza.

Tabla 13. Puntuaciones obtenidas en las evaluaciones antes (Pre) y después (Post) de la enseñanza de los algoritmos de soporte vital básico.

| Evaluación | Grupo | Pre | Post |
|-----------------|------------------------------|-----------|--------------|
| Test | <i>Grupo clase magistral</i> | 1.5 (1-2) | 2 (2-3) # |
| | <i>Grupo caso/problema</i> | 1 (1-2) | 2 (2-3) # |
| | <i>P (entre grupos)</i> | 0.71 | 0.75 |
| Grabación video | <i>Grupo clase magistral</i> | 1 (0-1) | 2 (2-3) # |
| | <i>Grupo caso/problema</i> | 1 (0-1) | 2 (1.75-3) # |
| | <i>P (entre grupos)</i> | 0.72 | 0.43 |

Los valores están expresados como medianas y percentiles 25-75. (#) $p < 0.001$ versus T1 intragrupo.

El análisis de los incrementos absolutos de puntuación y la ganancia de puntuación en el test y en el video no mostró ninguna diferencia significativa entre los grupos (Tabla 14).

Tabla 14. Incrementos absolutos de puntuación y ganancia de puntuación (Δ) en las evaluaciones.

| Puntuación | <i>Grupo clase magistral</i> (n = 34) | <i>Grupo caso/problema</i> (n = 34) | <i>P</i> |
|-----------------------------------|--|--|----------|
| Test post – Test pre | 1 (0-1) | 1 (0-1.25) | 0.52 |
| Δ (Test post – Test pre) | 50 (0-58.33) | 50 (0-83.33) | 0.31 |
| Video post – Video pre | 1.5 (1-2) | 1 (1-2) | 0.25 |
| Δ (Video post – Video pre) | 50 (25-66.66) | 50 (0-83.33) | 0.33 |

Pre: puntuación obtenida antes de la enseñanza; Post: puntuación obtenida después de la enseñanza.
 Δ Test = $100 * [(valor\ final - valor\ inicial) / (3 - valor\ inicial)]$. Δ Video = $100 * [(valor\ final - valor\ inicial) / (4 - valor\ inicial)]$. Datos expresados como mediana y percentiles 25-75.

No encontramos diferencias entre los grupos en el número de estudiantes que mejoraron sus puntuaciones después de la enseñanza (Tabla 15). En ambos grupos, los porcentajes de estudiantes que mejoraron sus puntuaciones fueron relevantes ($> 50\%$); los que mejoraron sus puntuaciones en la detección de errores (video) fue mayor que los que mejoraron sus puntuaciones en la toma las decisiones correctas (test). En el *grupo clase magistral* el número de estudiantes que mejoraron sus puntuaciones en el video fue significativamente mayor a los que mejoraron sus puntuaciones en el test.

Tabla 15. Número de estudiantes que obtuvieron puntuaciones mejores (rangos positivos), peores (rangos negativos) o iguales (rangos iguales) en las evaluaciones.

| Diferencias de puntuación | Rangos | Grupo clase magistral (n = 34) | Grupo caso/problema (n = 34) | P |
|---------------------------|----------|--------------------------------|------------------------------|------|
| Test T2 – Test T1 | Positivo | 19 (55.9) ^a | 20 (58.8) | 0.8 |
| | Negativo | 2 (5.9) | 4 (11.8) | 0.39 |
| | Igual | 13 (38.2) | 10 (29.4) | 0.44 |
| Video T2 – Video T1 | Positivo | 29 (85.3) ^a | 28 (82.3) | 0.74 |
| | Negativo | 2 (5.9) | 0 (0.0) | 0.15 |
| | Igual | 3 (8.8) | 6 (17.6) | 0.28 |

Test de Chi-cuadrado. Valores expresados como número de casos (porcentaje intragrupo). (a) $p = 0,012$ intragrupo

5.2. Resultados Artículo 2. Carrero E, Gomar C, Penzo W, Rull M. Comparison between lecture-based approach and case/problem-based learning discussion for teaching pre-anaesthetic assessment. *Eur J Anaesthesiol* 2007; 24: 1008-15.

Cincuenta y cuatro residentes completaron el estudio, 29 en el grupo *clase magistral* y 25 en el *grupo caso/problema*. Tres participantes en el *grupo caso/problema* fueron excluidos porque no completaron la evaluación después de la enseñanza (Figura 7). El porcentaje de participación en el estudio fue 70,12%.

La mediana [percentiles 25-75] de edad fue 25 [24-25,5] años en el *grupo clase magistral* y 25 [24,5-26,5] años en el *grupo caso/problema* ($P = 0,21$). No encontramos diferencias significativas en la distribución de la edad de los residentes por grupo y año de estudio (Tabla 16).

Tabla 16. Distribución de los residentes de Anestesiología por edad según el grupo y año de estudio.

| Grupo | Año de estudio | Edad |
|------------------------------|----------------|---------------|
| <i>Grupo clase magistral</i> | 1999 | 25 [24–25,5] |
| | 2000 | 25 [24–25,75] |
| <i>Grupo caso/problema</i> | 1999 | 25 [24–26] |
| | 2000 | 25 [25–27] |

Datos expresados como mediana y percentiles 25-75. Test de Kruskal Wallis; $P = 0,40$.

El estudio incluyó 25 mujeres y 29 hombres; 10 mujeres en el grupo *clase magistral* y 15 en el *grupo caso/problema*, ($P = 0,06$). Tampoco encontramos diferencias significativas en la distribución del sexo de los residentes por grupo y año de estudio (Tabla 17).

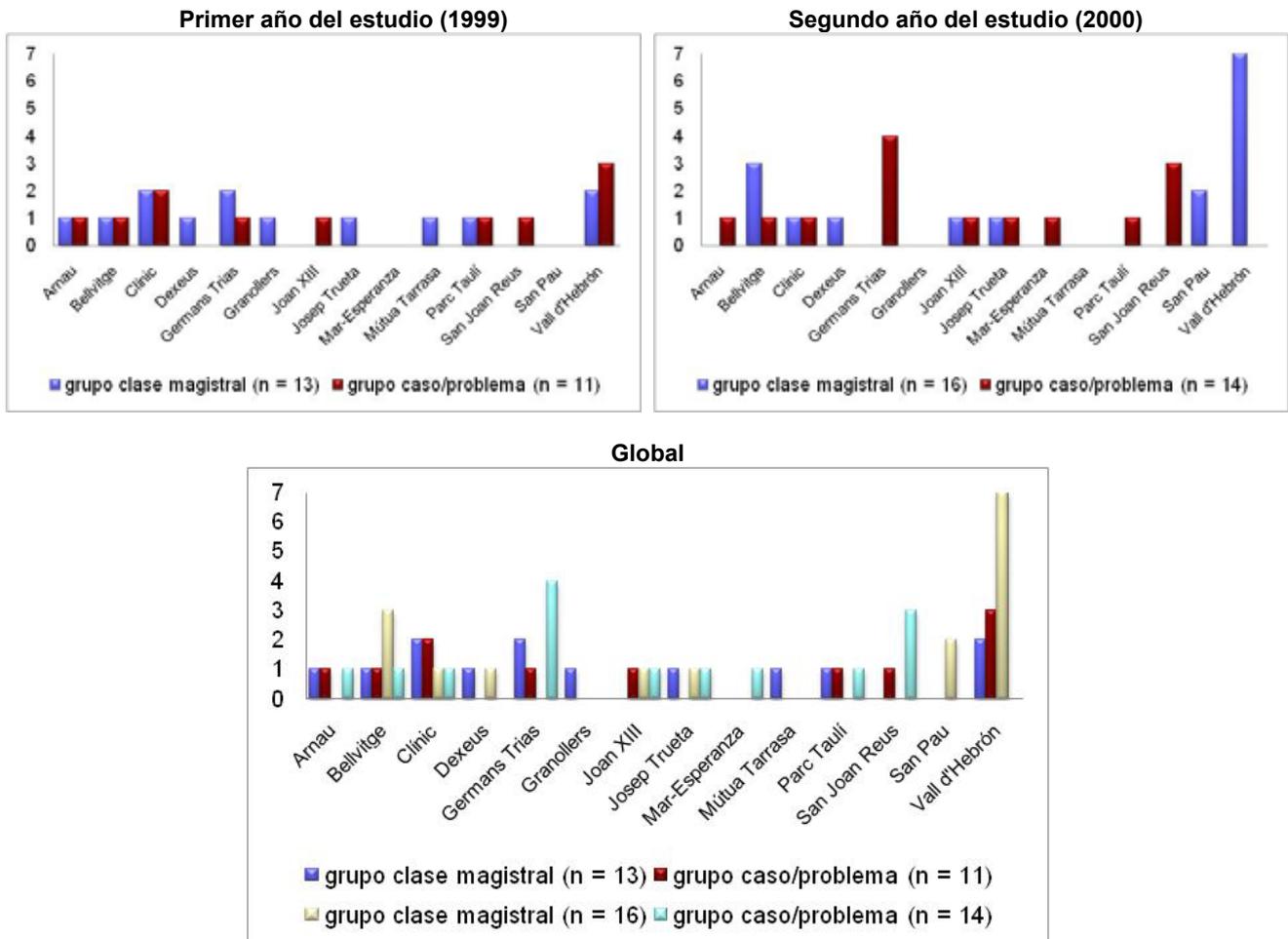
Tabla 17. Distribución de los residentes de Anestesiología por sexo según el grupo y año de estudio.

| Grupo | Año de estudio | Mujeres | Hombres | Total |
|------------------------------|----------------|----------|-----------|-------|
| <i>Grupo clase magistral</i> | 1999 | 5 (38,5) | 8 (61,5) | 13 |
| | 2000 | 5 (31,3) | 11 (68,8) | 16 |
| <i>Grupo caso/problema</i> | 1999 | 7 (63,3) | 4 (36,4) | 11 |
| | 2000 | 8 (57,1) | 6 (42,9) | 14 |

Datos expresados como número de casos (porcentaje). Test de Chi cuadrado; $P = 0,28$

Los randomización de los residentes por grupos resultó equilibrada en relación al hospital de procedencia, aunque hubo algunas excepciones: la mayor parte de los residentes de Bellvitge (4 de 6 residentes) y Vall d'Hebrón (9 de 12 residentes) quedaron ubicados en el *grupo clase magistral*; la

mayor parte de los residentes de Germans Trias i Pujol (5 de 7 residentes) y de Sant Joan de Reus (4 de 4 residentes) quedaron ubicados en el *grupo caso/problema* (Figura 13).

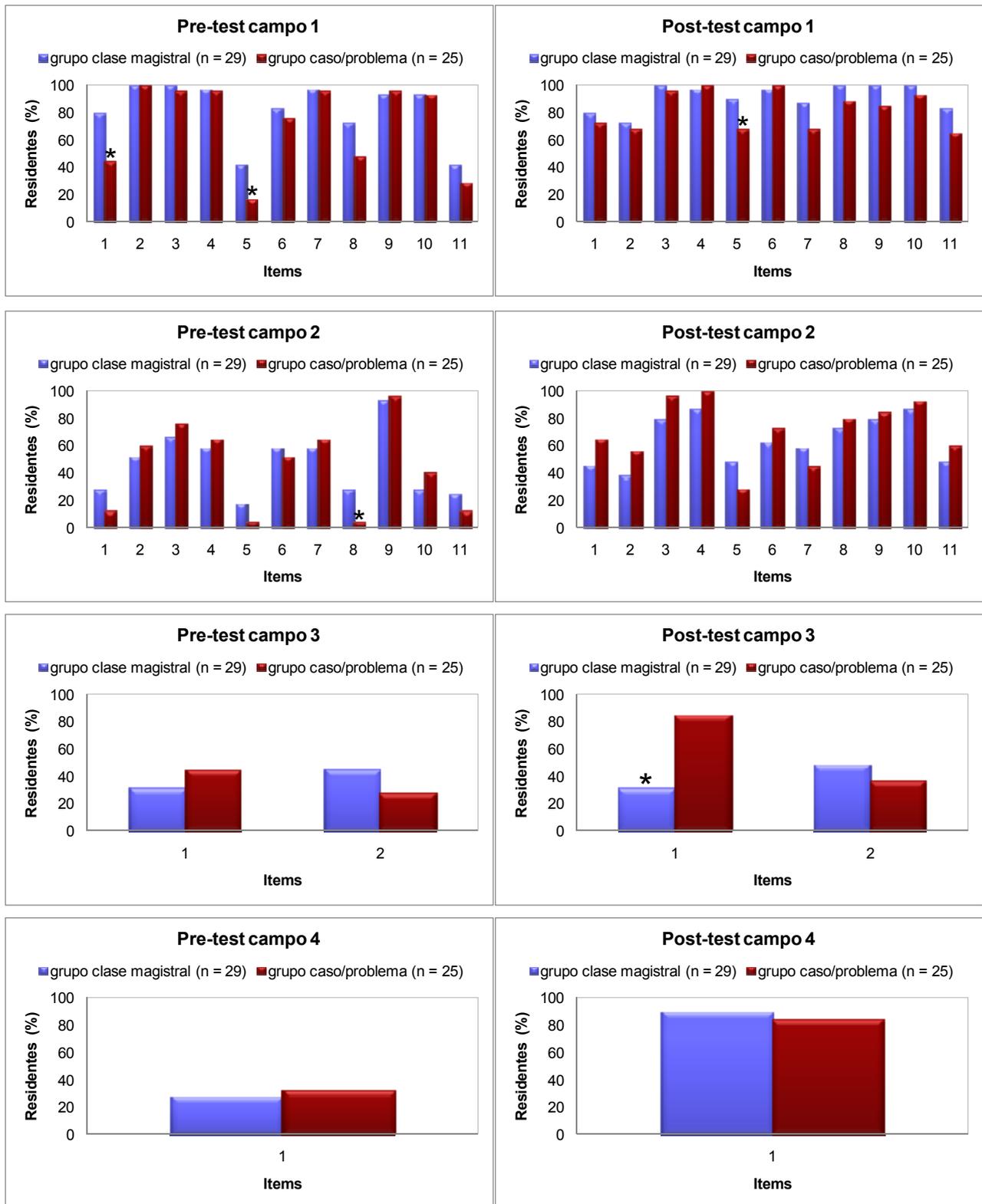


Las columnas representan el número de residentes en cada grupo que procedía de cada uno de los hospitales de Cataluña diferenciados por grupo y año de estudio.

Figura 13. Hospitales de procedencia de los residentes.

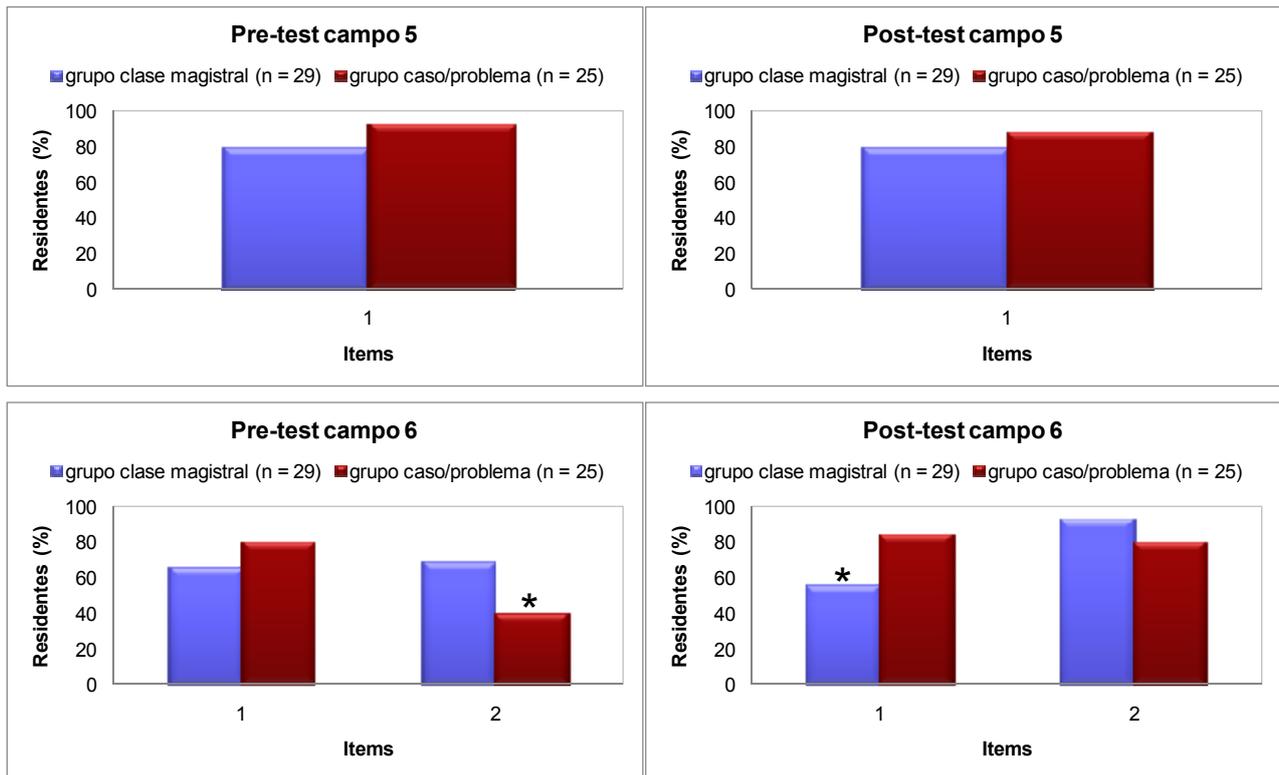
Antes de la enseñanza, encontramos diferencias significativas entre los grupos en el porcentaje de residentes que contestaron correctamente el ítem 1 “84 años” ($P = 0,007$) y el ítem 5 “hernia inguinal encarcelada” ($P = 0,042$) del campo 1 “Reconocimiento de datos clínicos con implicaciones anestésicas”; el ítem 8 “prednisona” ($P = 0,028$) del campo 2 “Razonamiento de datos clínicos con implicaciones anestésicas” y el ítem 2 “razonamiento de la elección de la técnica anestésica test 2” ($P = 0,033$) del campo 6 “Razonamiento de la elección técnica anestésica”. Después de la enseñanza, encontramos diferencias significativas entre los grupos en el ítem 5 “sustitución valvular aórtica” ($P = 0,049$) del campo 1; el ítem 1 “clase ASA II” ($P = 0,000$) del campo 3 “Clase ASA” y el ítem 1 “razonamiento de la elección de la técnica anestésica test 3” ($P = 0,049$) del campo 6 (Figuras 14A y 14B).

Resultados



Las columnas representan el porcentaje de residentes de cada grupo que contestó correctamente cada uno de los ítems del campo 1 "Reconocimiento de datos clínicos con implicaciones anestésicas", campo 2 "Razonamiento de datos clínicos con implicaciones anestésicas", campo 3 "Clase ASA" y campo 4 "Clase Mallampati". (*) Test Chi-Cuadrado, $P < 0.05$ entre grupos.

Figura 14A. Porcentaje de residentes que acertaron los ítems de los campos 1-4 medidos antes (pre-test) y después de la enseñanza (post-test).



Las columnas representan el porcentaje de residentes de cada grupo que contestó correctamente cada uno de los ítems del campo 5 “Elección de la técnica anestésica” y campo 6 “Razonamiento de la elección técnica anestésica”. (*) Test Chi-Cuadrado, $P < 0.05$ entre grupos.

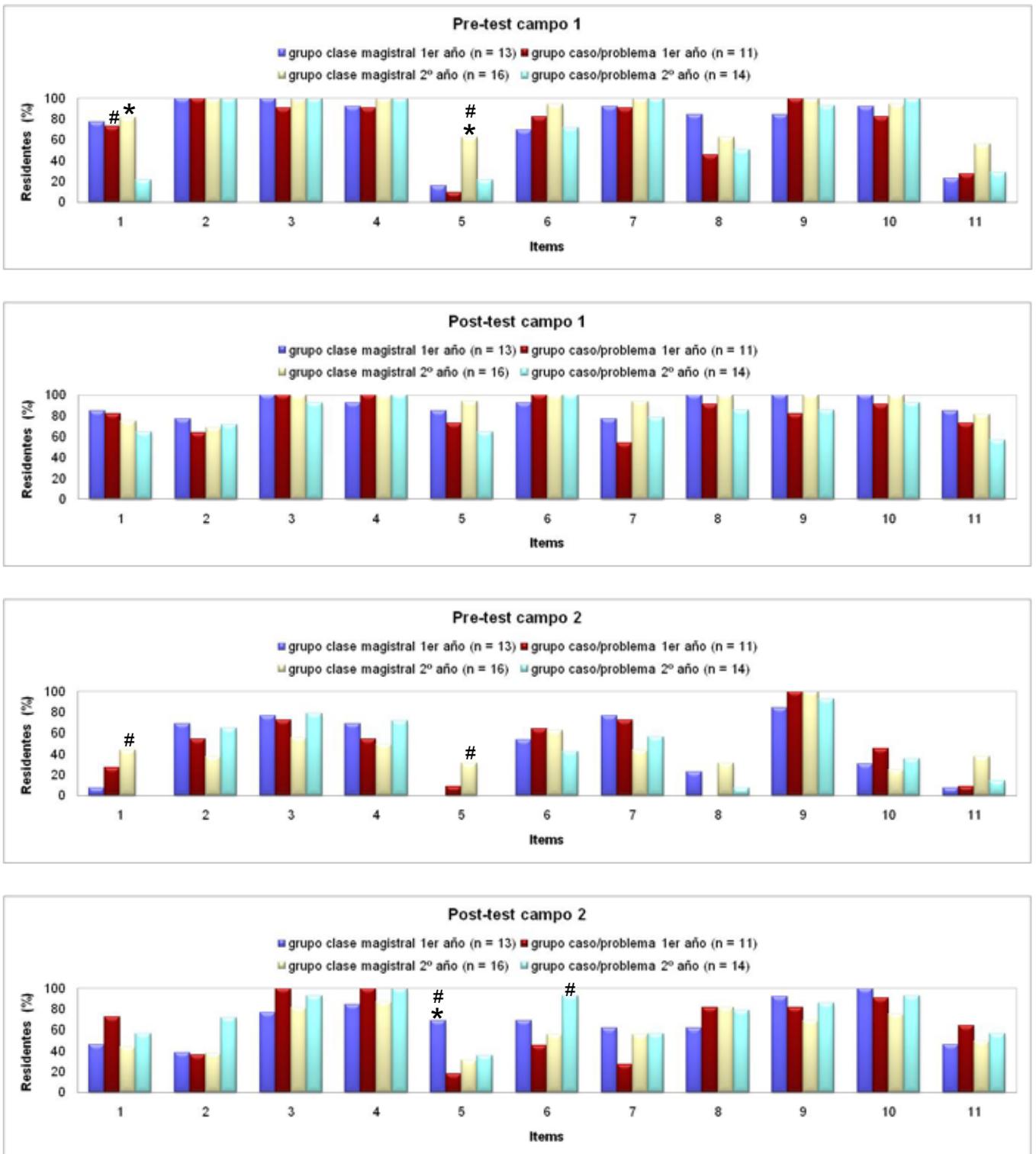
Figura 14B. Porcentaje de residentes que acertaron los ítems de los campos 5 y 6 medidos antes (pre-test) y después de la enseñanza (post-test).

El análisis por grupo y año de estudio mostró los siguientes hallazgos relevantes (Figuras 15A-15C):

Antes de la enseñanza, encontramos diferencias significativas entre los dos años de estudio en el porcentaje de aciertos en el ítem 1 “edad 84 años” ($P = 0,017$), grupo caso/problema y en el ítem 5 “hernia inguinal encarcelada” ($P = 0,01$), grupo clase magistral, del campo 1 “Reconocimiento de datos clínicos con implicaciones anestésicas”; en el ítem 1 “edad 84 años” ($P = 0,044$) y en el ítem 5 “hernia inguinal encarcelada” ($P = 0,048$), grupo clase magistral, del campo 2 “Razonamiento de datos clínicos con implicaciones anestésicas” y en el ítem 2 “clase ASA III” ($P = 0,017$), grupo clase magistral, del campo 3 “Clase ASA”. El segundo año del estudio, el porcentaje de residentes que contestó correctamente el ítem 1 y el ítem 5 del campo 1 fue significativamente mayor en el grupo clase magistral ($P = 0,001$ y $P = 0,024$ respectivamente).

Después de la enseñanza, encontramos diferencias significativas entre los dos años de estudio en el porcentaje de aciertos en el ítem 5 “sustitución valvular aórtica” ($P = 0,042$), grupo clase magistral, y en el ítem 6 “anticoagulantes orales” ($P = 0,021$), grupo caso/problema, del campo 2 “Razonamiento de datos clínicos con implicaciones anestésicas”. El primer año de estudio, el porcentaje de aciertos del ítem 5 del campo 2 fue significativamente mayor en el grupo clase magistral que en el grupo caso/problema ($P = 0,012$).

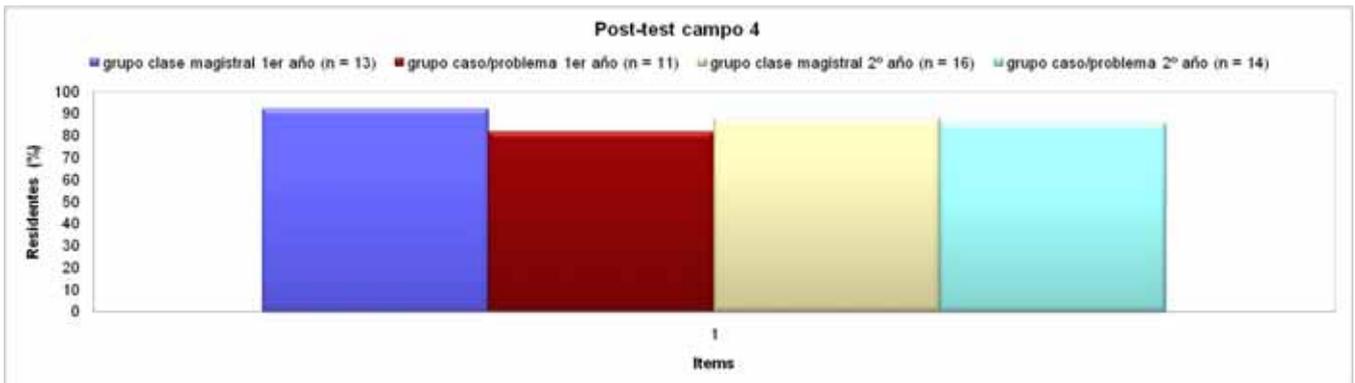
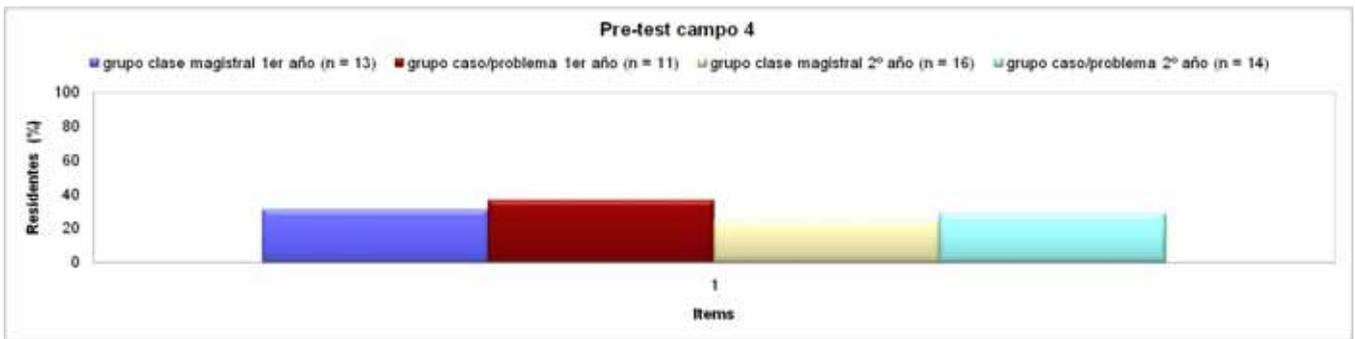
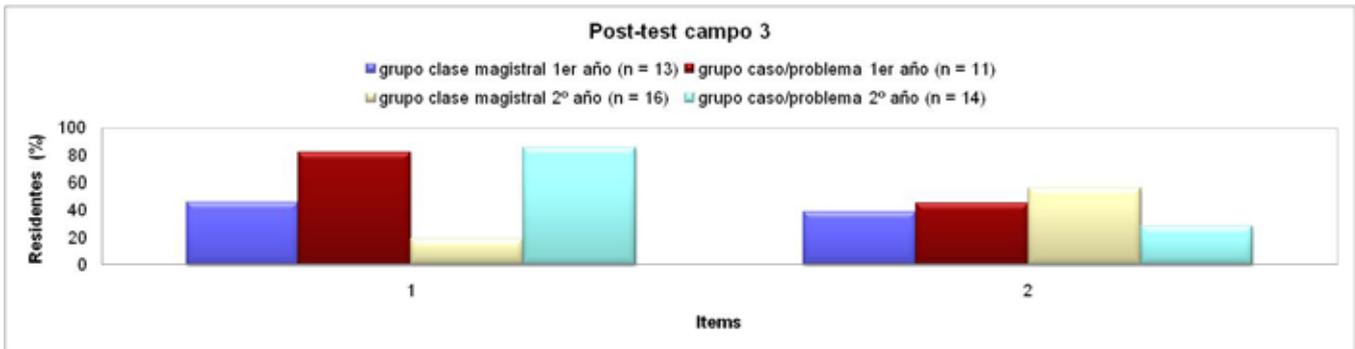
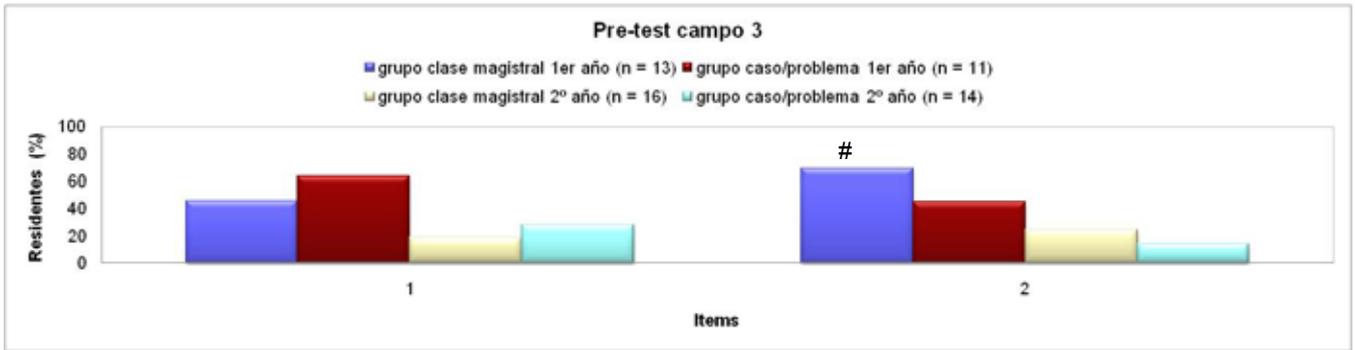
Resultados



Las columnas representan el porcentaje de residentes, por grupo y año de estudio, que contestó correctamente cada uno de los ítems del campo 1 “Reconocimiento de datos clínicos con implicaciones anestésicas” y campo 2 “Razonamiento de datos clínicos con implicaciones anestésicas”. (*) Test Chi-Cuadrado, $P < 0.05$ entre grupos el mismo año del estudio. (#) Test Chi-Cuadrado, $P < 0.05$ intragrupo entre los dos años del estudio.

Figura 15A. Porcentaje de residentes del primer año (1999) y segundo año (2000) del estudio que acertaron los ítems de los campos 1 y 2 medidos antes (pre-test) y después (post-test) de la enseñanza.

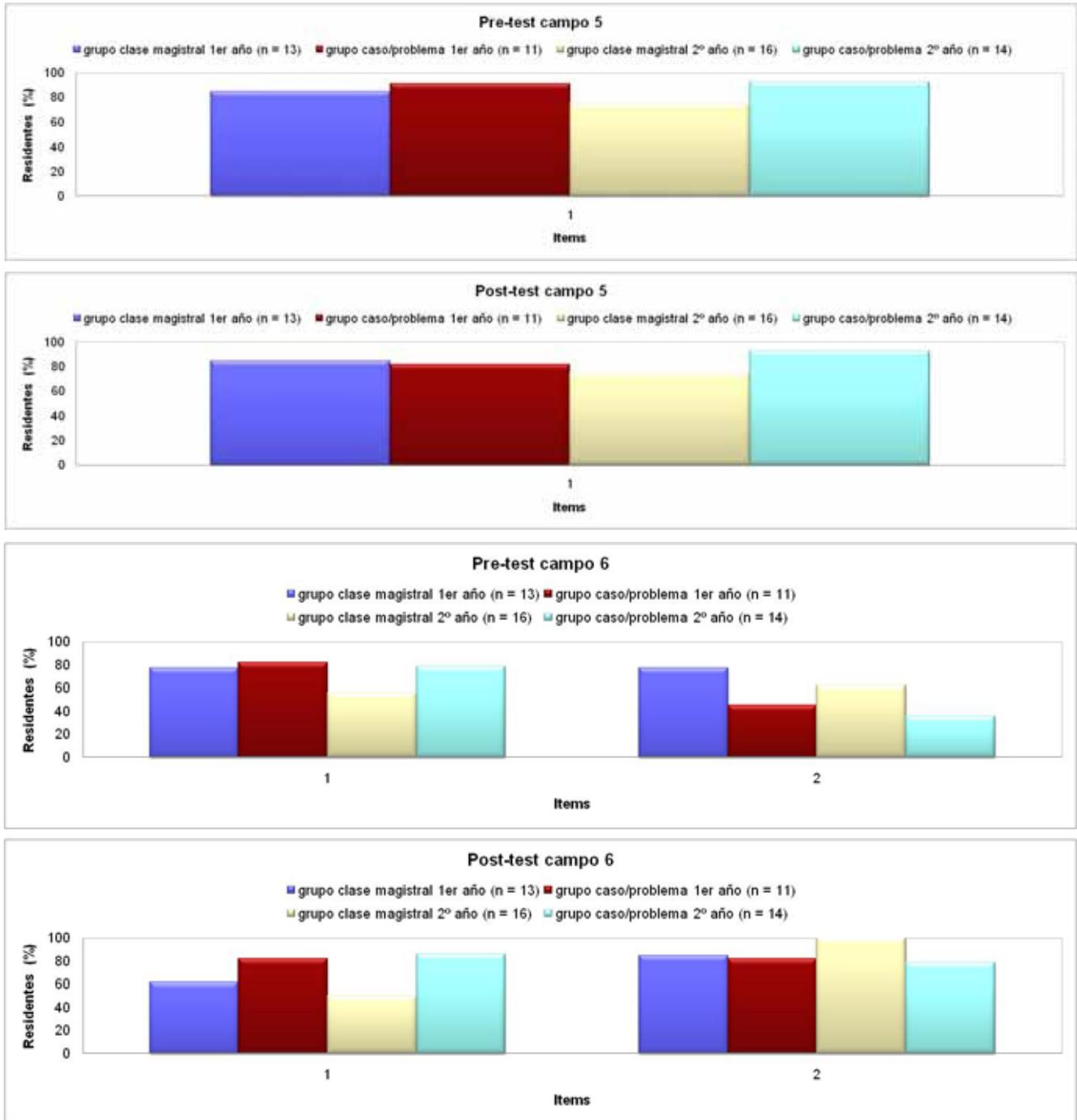
Resultados



Las columnas representan el porcentaje de residentes, por grupo y año de estudio, que contestó correctamente cada uno de los ítems del campo 3 “Clase ASA” y campo 4 “Clase Mallampati”. (*) Test *Chi-Cuadrado*, $P < 0.05$ entre grupos el mismo año del estudio. (#) Test *Chi-Cuadrado*, $P < 0.05$ intragrupo entre los dos años del estudio.

Figura 15B. Porcentaje de residentes del primer año (1999) y segundo año (2000) del estudio que acertaron los ítems de los campos 3 y 4 medidos antes (pre-test) y después (post-test) de la enseñanza.

Resultados



Las columnas representan el porcentaje de residentes, por grupo y año de estudio, que contestó correctamente cada uno de los ítems del campo 5 “Elección de la técnica anestésica” y campo 6 “Razonamiento de la elección de la técnica anestésica”. (*) Test *Chi-Cuadrado*, $P < 0.05$ entre grupos el mismo año del estudio. (#) Test *Chi-Cuadrado*, $P < 0.05$ intragrupo entre los dos años del estudio.

Figura 15C. Porcentaje de residentes del primer año (1999) y segundo año (2000) del estudio que acertaron los ítems de los campos 5 y 6 medidos antes (pre-test) y después (post-test) de la enseñanza.

Los resultados de las evaluaciones antes (*pre-tests*) y después de la enseñanza (*post-tests*) están resumidos en la Tabla 18. El índice *kappa* de concordancia entre evaluadores fue > 0.8 para todas las evaluaciones. Las puntuaciones obtenidas en los *pre-tests* fueron similares entre los grupos para todos los campos medidos excepto para el campo 1, que fue significativamente mayor en el *grupo clase magistral* ($P = 0,006$). Después de la enseñanza, las puntuaciones del campo 1, campo 2, y campo 4 mejoraron significativamente en el *grupo clase magistral* ($P = 0,001$, $P = 0,001$ y $P = 0,000$, respectivamente). En el *grupo caso/problema*, encontramos mejorías significativas de las puntuaciones en el campo 1 ($P = 0,001$), campo 2 ($P = 0,000$), campo 3 ($P = 0,042$), campo 4 ($P = 0,001$) y campo 6 ($P = 0,022$). La puntuación del campo 1 *post-test* fue significativamente mayor en el *grupo clase magistral* ($P = 0,005$), mientras que la puntuación del campo 3 *post-test* fue significativamente mayor en el *grupo caso/problema* ($P = 0,044$). No encontramos diferencias significativas entre los grupos para los otros campos cuando comparamos las puntuaciones *post-tests*.

Tabla 18. Puntuaciones obtenidas antes de la enseñanza (pre-test) y después de la enseñanza (post-test) de acuerdo a los campos de conocimiento medidos.

| Campo | Test | Grupo clase magistral (n = 29) | Grupo caso/problema (n = 25) |
|--|------|--------------------------------|------------------------------|
| Campo 1 "Reconocimiento de datos clínicos con implicaciones anestésicas" | Pre | 9 (8-10)* | 8 (7-9) |
| | Post | 10 (9.5-11)* # | 10 (8-10) # |
| Campo 2 "Razonamiento de datos clínicos con implicaciones anestésicas" | Pre | 5 (4-6) | 4 (4-6) |
| | Post | 7 (6-9)# | 8 (7-9)# |
| Campo 3 "Clase ASA" | Pre | 1 (0-1,5) | 1 (0-1) |
| | Post | 1 (0-1) | 1 (1-2)* # |
| Campo 4 "Clase Mallampati" | Pre | 0 (0-1) | 0 (0-1) |
| | Post | 1 (1-1)# | 1 (1-1)# |
| Campo 5 "Elección de la técnica anestésica" | Pre | 1 (1-1) | 1 (1-1) |
| | Post | 1 (1-1) | 1 (1-1) |
| Campo 6 "Razonamiento de la elección de la técnica anestésica" | Pre | 1 (1-2) | 1 (1-2) |
| | Post | 2 (1-2) | 2 (1-2)# |

Datos expresados como mediana y percentiles 25-75. (*) Test *U* de Mann-Whitney, $P < 0.05$ entre grupos.

(#) Test de los rangos de Wilcoxon, $P < 0.05$ intragrupo. (n) = número de casos analizados.

La Tabla 19 muestra los resultados de las evaluaciones por grupo y año de estudio. En el *grupo clase magistral*, la puntuación del campo 1 antes de la enseñanza fue mayor el segundo año del estudio comparada con el primer año del estudio ($P = 0,028$) y la puntuación del campo 3 mayor el primer año ($P = 0,028$). Comparada con el *grupo caso/problema* la puntuación del campo 1 antes de la enseñanza el segundo año del estudio fue significativamente mayor en el *grupo clase magistral* ($P = 0,004$); después de la enseñanza, dicha diferencia se mantuvo significativa ($P = 0,043$). El primer año del estudio, las puntuaciones del campo 1, campo 2 y campo 4 mejoraron significativamente en el *grupo clase magistral* ($P = 0,007$, $P = 0,01$ y $P = 0,011$, respectivamente). En el *grupo caso/problema*, encontramos mejorías significativas en el campo 1 ($P = 0,018$) y campo 2 ($P = 0,024$). El segundo año del estudio, las puntuaciones del campo 2 y campo 4 repitieron la mejoría significativa en el *grupo clase magistral* ($P = 0,025$, $P = 0,002$, respectivamente). En el *grupo caso/problema*, encontramos mejorías significativas en el campo 1 ($P = 0,02$), campo 2 ($P = 0,002$), campo 3 ($P = 0,02$), campo 4 ($P = 0,005$) y campo 6 ($P = 0,035$).

Tabla 19. Puntuaciones obtenidas en cada uno de los dos años del estudio de acuerdo a los campos de conocimiento medidos.

| Campo | Test | Grupo clase magistral | | Grupo caso/problema | |
|---------|------|----------------------------|---------------------------|----------------------------|---------------------------|
| | | 1 ^{er} año (n=13) | 2 ^o año (n=16) | 1 ^{er} año (n=11) | 2 ^o año (n=14) |
| Campo 1 | Pre | 8 (7.5 – 9) | 9.5 (8.25 – 11) *, & | 8 (8 – 8) | 7.5 (7 – 9) |
| | Post | 11 (9.5 – 11) # | 10 (9.25 – 11) * | 10 (8 – 10) # | 9.5 (7.75 – 10) # |
| Campo 2 | Pre | 5 (4 – 6) | 5 (3.25 – 7) | 6 (3 – 7) | 4 (4 – 5.25) |
| | Post | 7 (6 – 9) # | 7.5 (6 – 9) # | 7 (6 – 8) # | 8.5 (7 – 9.25) # |
| Campo 3 | Pre | 1 (0.5 – 2) | 0 (0 – 1) & | 1 (0 – 2) | 0 (0 – 1) |
| | Post | 1 (0 – 1) | 1 (0 – 1) | 1 (1 – 2) | 1 (1 – 2) # |
| Campo 4 | Pre | 0 (0 – 1) | 0 (0 – 0.75) | 0 (0 – 1) | 0 (0 – 1) |
| | Post | 1 (1 – 1) # | 1 (1 – 1) # | 1 (1 – 1) | 1 (1 – 1) # |
| Campo 5 | Pre | 1 (1 – 1) | 1 (0.25 – 1) | 1 (1 – 1) | 1 (1 – 1) |
| | Post | 1 (1 – 1) | 1 (0.25 – 1) | 1 (1 – 1) | 1 (1 – 1) |
| Campo 6 | Pre | 2 (1 – 2) | 1 (1 – 2) | 1 (1 – 2) | 1 (1 – 2) |
| | Post | 2 (1 – 2) | 1.5 (1 – 2) | 2 (1 – 2) | 2 (1 – 2) # |

Pre-test: evaluación antes de la enseñanza; *Post-test*: evaluación después de la enseñanza. Campo 1: "Reconocimiento de datos clínicos con implicaciones anestésicas"; Campo 2: "Razonamiento de datos clínicos con implicaciones anestésicas"; Campo 3: "Clase ASA"; Campo 4: "Clase Mallampati"; Campo 5: "Elección de la técnica anestésica"; Campo 6: "Razonamiento de la elección de la técnica anestésica". Datos expresados como mediana y percentiles 25-75. (*) Test U de Mann-Whitney, $P < 0.05$ entre grupos para el mismo año de estudio y el mismo test. (#) Test de los rangos de Wilcoxon o test de McNemar para variables dicotómicas, $P < 0.05$ intragrupo comparando el post-test con el pre-test para el mismo año de estudio. (&) Test U de Mann-Whitney, $P < 0.05$ intragrupo comparando el primer año de estudio (1999) con el segundo año de estudio (2000) para el mismo test. (n) = número de casos analizados.

El análisis de las diferencias entre las puntuaciones del *post-test* y el *pre-test* no reveló diferencias significativas entre los dos grupos para ninguno de los campos medidos (Tabla 20). Tampoco encontramos diferencias significativas entre los 4 grupos por año de estudio para ninguno de los campos medidos (Tabla 21).

Tabla 20. Incrementos absolutos de puntuación para cada campo del conocimiento medido.

| Post-test menos pre-test | Grupo clase magistral (n = 29) | Grupo caso/problema (n = 25) | P |
|----------------------------|-----------------------------------|---------------------------------|------|
| Campo 1 post – Campo 1 pre | 1 [0-2] | 1 [0-2] | 0,72 |
| Campo 2 post – Campo 2 pre | 2 (0-4] | 3 [1-5] | 0,15 |
| Campo 3 post – Campo 3 pre | 0 [(-1)-1] | 0 [0-1] | 0,11 |
| Campo 4 post – Campo 4 pre | 1 [0-1] | 1 [0-1] | 0,49 |
| Campo 5 post – Campo 5 pre | 0 [0-0] | 0 [0-0] | 0,78 |
| Campo 6 post – Campo 6 pre | 0 [(-1)-1] | 0 [0-1] | 0,25 |

Pre-test: evaluación antes de la enseñanza; *Post-test*: evaluación después de la enseñanza. Campo 1: "Reconocimiento de datos clínicos con implicaciones anestésicas"; Campo 2: "Razonamiento de datos clínicos con implicaciones anestésicas"; Campo 3: "Clase ASA"; Campo 4: "Clase Mallampati"; Campo 5: "Elección de la técnica anestésica"; Campo 6: "Razonamiento de la elección de la técnica anestésica". Datos expresados como mediana y percentiles 25-75. Test U de Mann-Whitney (diferencias entre los 2 grupos). (n) = número de casos analizados.

Tabla 21. Incrementos absolutos de puntuación para cada campo del conocimiento medido en cada uno de los dos años del estudio.

| Post-test menos pre-test | Grupo clase magistral | | Grupo caso/problema | | P |
|----------------------------|-----------------------|--------------------|---------------------|--------------------|------|
| | 1er año (n = 13) | 2º año (n = 16) | 1er año (n = 11) | 2º año (n = 14) | |
| Campo 1 post – Campo 1 pre | 2 [0,5-3] | 0 [0-1] | 2 [0-2] | 1 [0,75-2] | 0,22 |
| Campo 2 post – Campo 2 pre | 3 [0-5] | 1,5 [0,25-3,75] | 1 [1-5] | 4,5 [1,75-5] | 0,19 |
| Campo 3 post – Campo 3 pre | 0 [(-1)-0] | 0 [0-1] | 0 [(-1)-1] | 1 [0-1,25] | 0,09 |
| Campo 4 post – Campo 4 pre | 1 [0-1] | 1 [0-1] | 1 [0-1] | 1 [0-1] | 0,89 |
| Campo 5 post – Campo 5 pre | 0 [0-0] | 0 [0-0] | 0 [0-0] | 0 [0-0] | 0,96 |
| Campo 6 post – Campo 6 pre | 0 [(-1)-1] | 0 [0-1] | 0 [0-1] | 0,5 [0-1] | 0,40 |

Datos expresados como mediana y percentiles 25-75. Test Kruskal Wallis (diferencias entre los 4 grupos por año). (n) = número de casos analizados. Primer año de estudio: 1999; 2º año: 2000.

No encontramos diferencias entre los grupos en el número de residentes que mejoraron sus puntuaciones después de la enseñanza (Tabla 22). Los porcentajes de residentes que mejoraron sus puntuaciones en el campo 1, campo 2 y campo 4 fueron relevantes (> 50%) en ambos grupos. Encontramos diferencias significativas en ambos grupos al comparar el porcentaje de residentes que mejoró su puntuación entre los distintos campos de conocimiento medidos ($P = 0,000$ y $P = 0,000$, respectivamente).

Tabla 22. Número de residentes que obtuvieron puntuaciones mejores (rangos positivos), peores (rangos negativos) o iguales (rangos iguales) en las evaluaciones.

| Diferencias de puntuación | Rangos | Grupo clase/magistral (n = 29) | Grupo caso/problema (n = 25) | P |
|----------------------------|----------|--------------------------------|------------------------------|------|
| Campo 1 post – Campo 1 pre | Positivo | 17 (58,6) | 18 (72) | 0,30 |
| | Negativo | 4 (13,8) | 3 (12,0) | 1,00 |
| | Igual | 8 (27,6) | 4 (16,0) | 0,30 |
| Campo 2 post – Campo 2 pre | Positivo | 20 (69,0) [#] | 21 (84) [#] | 0,19 |
| | Negativo | 3 (10,3) | 1 (4,0) | 0,61 |
| | Igual | 6 (20,7) | 3 (12,0) | 0,48 |
| Campo 3 post – Campo 3 pre | Positivo | 8 (27,6) | 12 (48,0) | 0,12 |
| | Negativo | 8 (27,6) | 4 (16,0) | 0,30 |
| | Igual | 13 (44,8) | 9 (36,0) | 0,51 |
| Campo 4 post – Campo 4 pre | Positivo | 19 (65,5) | 14 (56,0) | 0,47 |
| | Negativo | 1 (3,4) | 1 (4,0) | 1,00 |
| | Igual | 9 (31,0) | 10 (40,0) | 0,49 |
| Campo 5 post – Campo 5 pre | Positivo | 5 (17,2) | 1 (4,0) | 0,20 |
| | Negativo | 5 (17,2) | 2 (8,0) | 0,43 |
| | Igual | 19 (65,5) | 22 (88,0) | 0,05 |
| Campo 6 post – Campo 6 pre | Positivo | 11 (37,9) | 11 (44,0) | 0,65 |
| | Negativo | 8 (27,6) | 3 (12,0) | 0,15 |
| | Igual | 10 (34,5) | 11 (44,0) | 0,47 |

Test de *Chi-cuadrado* (diferencias entre grupos). (#) Test de Q de Cochran (diferencias intragrupo para rangos positivos). Valores expresados como número de casos (porcentaje intragrupo).

Tampoco encontramos diferencias entre los 4 grupos por año de estudio en el número de residentes que mejoraron sus puntuaciones después de la enseñanza (Tabla 23). Los porcentajes de residentes que mejoraron sus puntuaciones en el campo 2 y campo 4 fueron relevantes (> 50%) en los 4 grupos por año de estudio. Encontramos diferencias significativas en los 4 grupos por año de estudio al comparar el porcentaje de residentes que mejoró su puntuación entre los distintos campos de conocimiento medidos (*grupo clase magistral*, primer año de estudio $P = 0,001$; *grupo clase magistral*, segundo año del estudio $P = 0,034$; *grupo caso/problema*, primer año del estudio; $P = 0,013$ y *grupo caso/problema* segundo año del estudio $P = 0,000$).

Tabla 23. Número de residentes que obtuvieron puntuaciones mejores (rangos positivos), peores (rangos negativos) o iguales (rangos iguales) en las evaluaciones cada uno de los años del estudio.

| Diferencias de puntuación | Rangos | Grupo clase/magistral | | Grupo caso/problema | | P |
|---------------------------|----------|------------------------|----------------------|-----------------------|------------------------|------|
| | | 1er año (n = 13) | 2º año (n = 16) | 1er año (n = 11) | 2º año (n = 14) | |
| Campo 1 post – | Positivo | 10 (76,9) [#] | 7 (43,8) | 7 (63,6) | 11 (78,6) | 0,17 |
| Campo 1 pre | Negativo | 2 (15,4) | 2 (12,5) | 1 (9,1) | 2 (14,3) | 0,97 |
| | Igual | 1 (7,7) | 7 (43,8) | 3 (27,3) | 1 (7,1) | 0,05 |
| Campo 2 post – | Positivo | 8 (61,5) | 12 (75) [#] | 9 (81,8) [#] | 12 (85,7) [#] | 0,49 |
| Campo 2 pre | Negativo | 1 (7,7) | 2 (12,5) | 1 (9,1) | 0 (0,0) | 0,62 |
| | Igual | 4 (30,8) | 2 (12,5) | 1 (9,1) | 2 (14,3) | 0,47 |
| Campo 3 post – | Positivo | 2 (15,4) | 6 (37,5) | 4 (36,4) | 8 (57,1) | 0,17 |
| Campo 3 pre | Negativo | 5 (38,5) | 3 (18,8) | 3 (27,3) | 1 (7,1) | 0,25 |
| | Igual | 6 (46,2) | 7 (43,8) | 4 (36,4) | 5 (35,7) | 0,93 |
| Campo 4 post – | Positivo | 9 (69,2) | 10 (62,5) | 6 (54,5) | 8 (57,1) | 0,88 |
| Campo 4 pre | Negativo | 1 (7,7) | 0 (0,0) | 1 (9,1) | 0 (0,0) | 0,46 |
| | Igual | 3 (23,1) | 6 (37,5) | 4 (36,4) | 6 (42,9) | 0,74 |
| Campo 5 post – | Positivo | 2 (15,4) | 3 (18,8) | 1 (9,1) | 0 (0,0) | 0,40 |
| Campo 5 pre | Negativo | 2 (15,4) | 3 (18,8) | 2 (18,2) | 0 (0,0) | 0,41 |
| | Igual | 9 (69,2) | 10 (62,5) | 8 (72,7) | 14 (100,0) | 0,09 |
| Campo 6 post – | Positivo | 4 (30,8) | 7 (43,8) | 4 (36,4) | 7 (50,0) | 0,76 |
| Campo 6 pre | Negativo | 5 (38,5) | 3 (18,8) | 2 (18,2) | 1 (7,1) | 0,24 |
| | Igual | 4 (30,8) | 6 (37,5) | 5 (45,5) | 6 (42,9) | 0,88 |

Test de Kruskal Wallis (diferencias entre los 4 grupos por año del estudio). (#) Test de Q de Cochran (diferencias intragrupo y año de estudio para rangos positivos). Valores expresados como número de casos (porcentaje intragrupo). Primer año de estudio: 1999; 2º año: 2000.

Encontramos correlación significativa entre el campo 1 “Reconocimiento de datos clínicos con implicaciones anestésicas” y el campo 2 “Razonamiento de datos clínicos con implicaciones anestésicas”, después de la enseñanza, en el grupo clase magistral ($P = 0,027$) y en grupo caso/problema ($P = 0,000$). El grado de correlación fue leve ($r^2 = 0,168$) en el grupo clase magistral y moderada ($r^2 = 0,467$) en el grupo caso/problema (Figura 16).

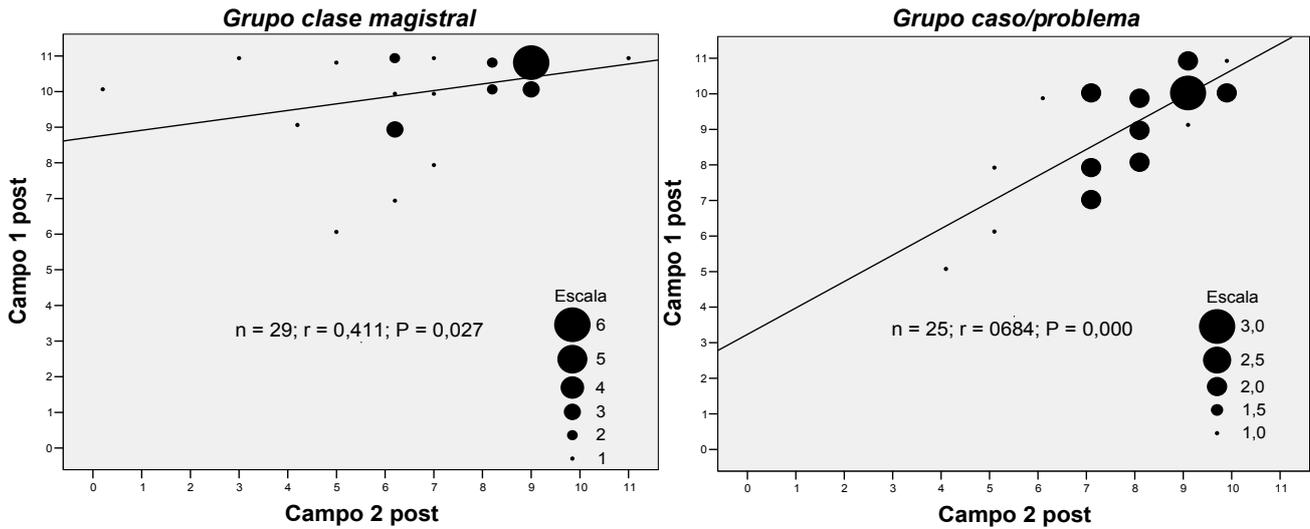


Diagrama de dispersión. Los puntos representan número de casos según la escala adjunta.

Figura 16. Correlación entre el campo 1 “Reconocimiento de datos clínicos con implicaciones anestésicas” y el campo 2 “Razonamiento de datos clínicos con implicaciones anestésicas” después de la enseñanza (post). (r) = coeficiente de correlación de Spearman.

El análisis por grupo y año de estudio mostró correlación lineal significativa entre el campo 1 y el campo 2 el segundo año del estudio, antes de la enseñanza en el *grupo clase magistral* ($P = 0,037$) y después de la enseñanza en el *grupo caso/problema* ($P = 0,000$). El grado de correlación fue leve ($r^2 = 0,275$) en el *grupo clase magistral* y fuerte ($r^2 = 0,725$) en el *grupo caso/problema* (Figura 17).

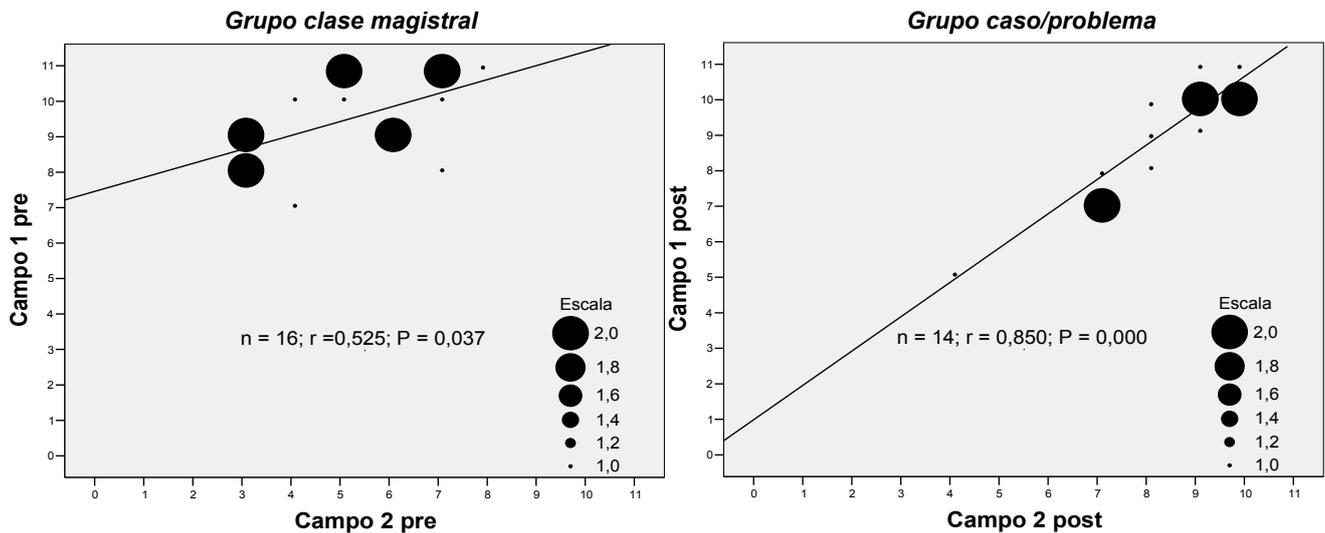


Diagrama de dispersión. Los puntos representan número de casos según la escala adjunta.

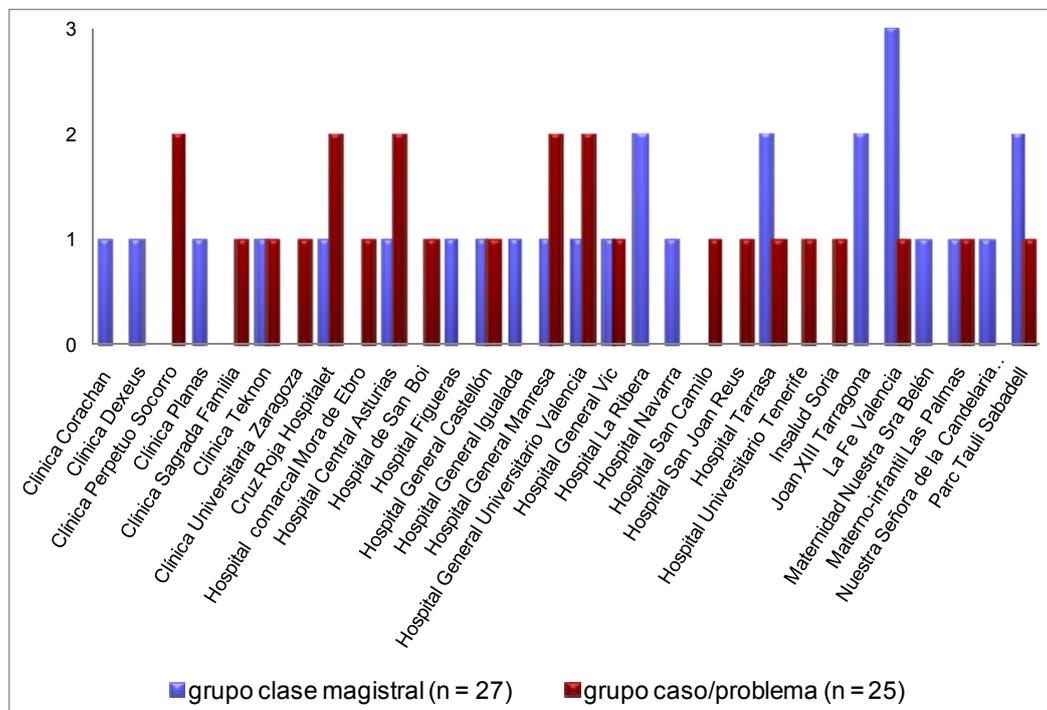
Figura 17. Correlación entre el campo 1 “Reconocimiento de datos clínicos con implicaciones anestésicas” y el campo 2 “Razonamiento de datos clínicos con implicaciones anestésicas” el segundo año del estudio (2000). (r) = coeficiente de correlación de Spearman. Pre: antes de la enseñanza, post = después de la enseñanza.

5.3. Resultados Artículo 3. Carrero EJ, Gomar C, Fábregas N, Penzo W, Castillo J, Villalonga A. Clase magistral versus aprendizaje basado en caso/problema para la enseñanza del embolismo aéreo en formación médica continuada. Rev Esp Anestesiol Reanim 2008; 55: 202-9.

Cincuenta y dos participantes se incluyeron finalmente en el estudio, 27 en el *grupo clase magistral* y 25 en el *grupo caso/problema* (Figura 9). Excluimos 7 participantes de los 59 iniciales por no cumplimentar o entregar alguno de los tests, 3 en el *grupo clase magistral* y 4 en el *grupo caso/problema*. El porcentaje de participación en el estudio fue 88,13%.

La distribución de los grupos por sexos resultó homogénea: 15 mujeres (55,6%) en el *grupo clase magistral* y 16 (64%) en el *grupo caso/problema* ($P = 0,53$). No encontramos diferencias significativas entre los grupos respecto al hospital de procedencia de los participantes (Figura 18), la comunidad autónoma del hospital (Figura 19) o la existencia de Servicio de Neurocirugía en el hospital (Figura 20). Los participantes en el estudio ejercían su actividad asistencial principalmente en hospitales de Cataluña (*grupo clase magistral*: 55,6%, *grupo caso/problema*: 60%; test *Chi cuadrado*, $P = 0,74$) y en segundo lugar en hospitales de Levante (*grupo clase magistral*: 25,9%, *grupo caso/problema*: 16%; $P = 0,38$). El 48,1% de los hospitales de procedencia en el *grupo clase magistral* y el 44% en el *grupo caso/problema* tenían Servicio de Neurocirugía ($P = 0,76$).

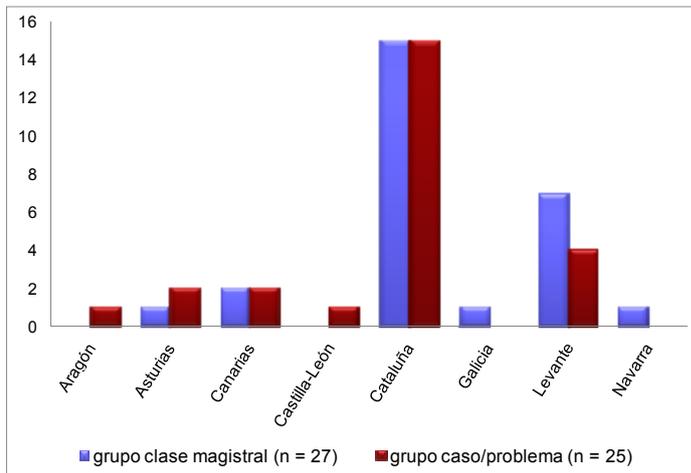
En cuanto a los años de experiencia profesional solo contestaron 6 participantes en cada grupo; las medias \pm DS expresadas en años fueron: *grupo clase magistral*: $11,17 \pm 4,83$ y *grupo caso/problema*: $9,83 \pm 4,53$; test T para 2 muestras independientes, $P = 0,63$.



Las columnas representan número de participantes.

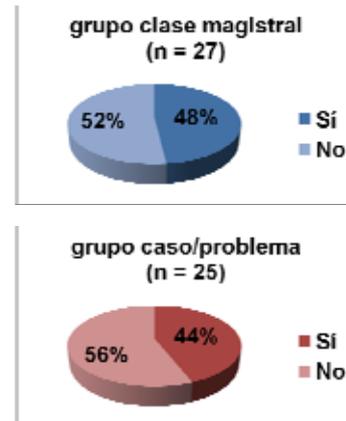
Figura 18. Hospitales de procedencia de los participantes en el estudio.

Resultados



Las columnas representan número de participantes.

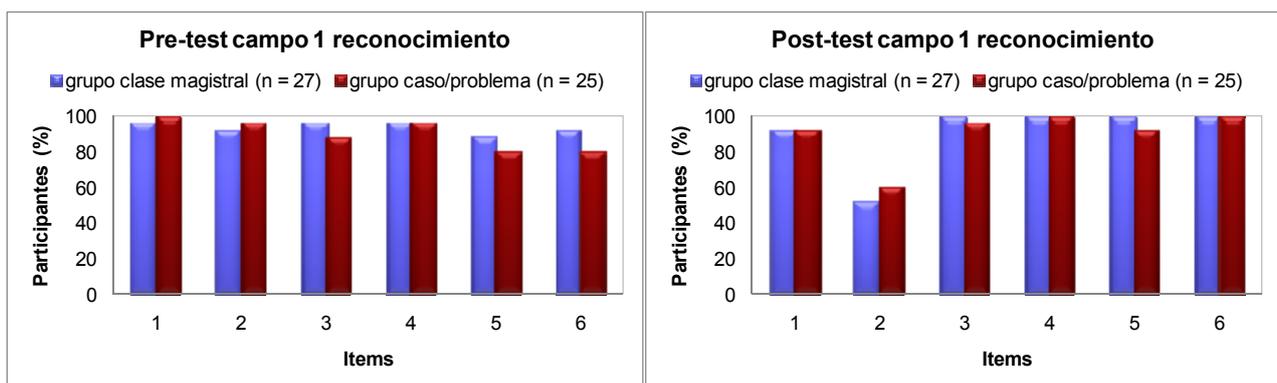
Figura 19. Comunidades autónomas donde ejercían su actividad laboral los participantes del estudio



Las áreas representan porcentaje de participantes.

Figura 20. Presencia de servicio de neurocirugía en el hospital de origen.

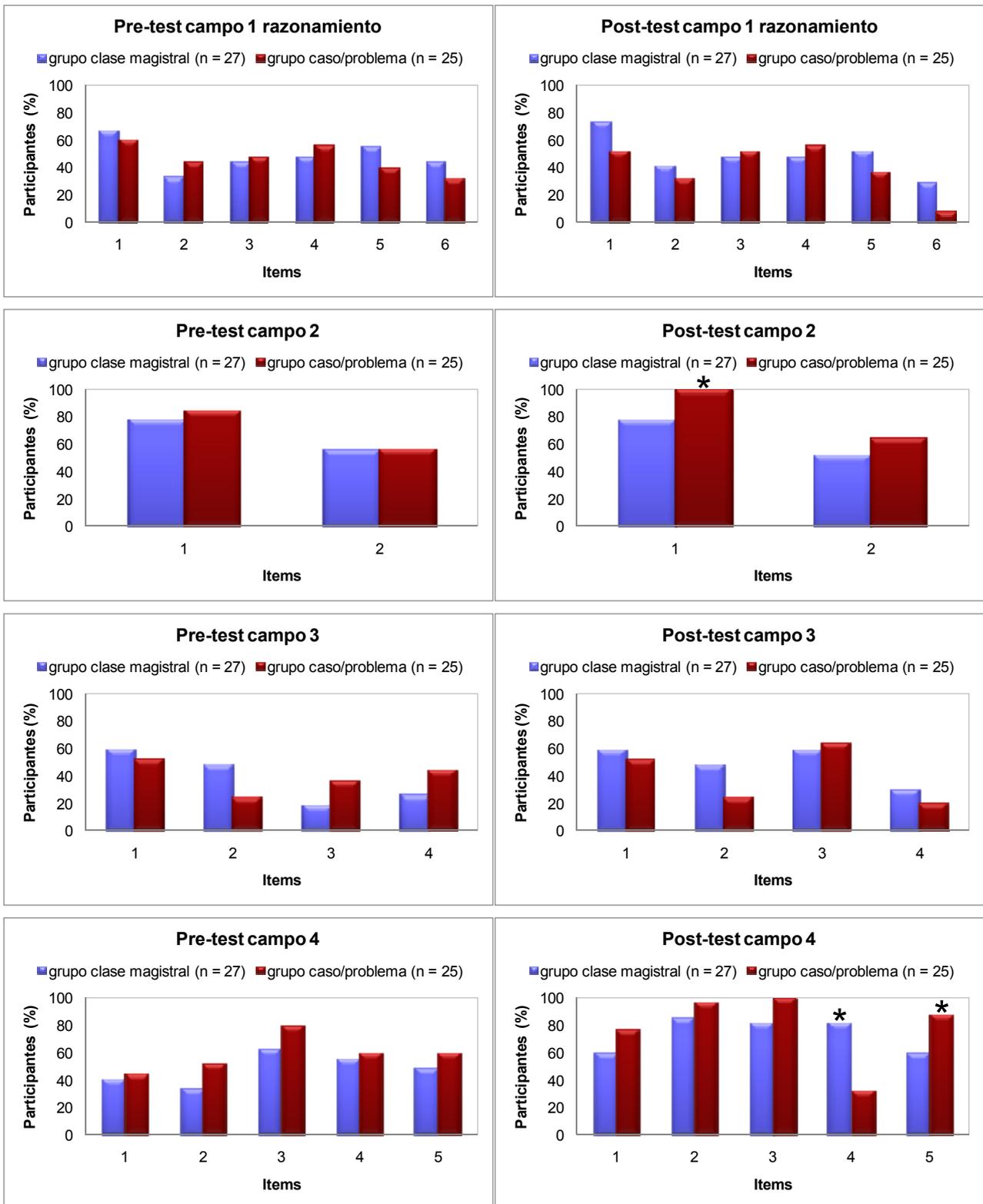
Antes de la enseñanza no encontramos diferencias significativas entre los grupos en ninguno de los ítems de los 4 campos medidos. Después de la enseñanza, el porcentaje de participantes que contestó correctamente el ítem 1 “*incluir embolismo aéreo en el diagnóstico diferencial*” del campo 2 “*Diagnóstico*” ($P = 0,02$) y el ítem 5 “*cambiar posición del paciente*” del campo 4 “*Tratamiento*” ($P = 0,02$) fue significativamente mayor en el *grupo caso/problema*; el porcentaje de participantes que contestó correctamente el ítem 4 “*aspirar aire por catéter venoso central*” del campo 4 “*Tratamiento*” fueron significativamente mayor en el *grupo clase magistral* ($P = 0,000$), (Figuras 21A y 21B).



Las columnas representan el porcentaje de participantes de cada grupo que contestó correctamente cada uno de los ítems.

Figura 21A. Porcentaje de participantes que reconocieron correctamente cada uno de los ítems del campo 1 “Factores de riesgo y síntomas” del tema embolismo aéreo en anestesia antes (pre-test) y después de la enseñanza (post-test).

Resultados



Las columnas representan el porcentaje de participantes de cada grupo que contestó correctamente cada uno de los ítems. (*) Test Chi-Cuadrado, $P < 0.05$ entre grupos.

Figura 21B. Porcentaje de participantes que razonaron correctamente cada uno de los ítems del campo 1 “Factores de riesgo y síntomas” y contestaron correctamente cada uno de los ítems de los campos 2 “Diagnóstico”, campo 3 “Monitorización” y campo 4 “Tratamiento” del tema embolismo aéreo en anestesia antes y después de la enseñanza.

Los resultados de las evaluaciones antes y después de la enseñanza están resumidos en la Tabla 24. El índice *kappa* de concordancia entre evaluadores fue > 0.8 para todas las evaluaciones.

Tabla 24. Puntuaciones en las evaluaciones antes (pretest) y después de la enseñanza (post-test) para los campos del conocimiento evaluados.

| Campo | Test | Grupo clase magistral (n=27) | Grupo caso/problema (n= 25) | P |
|--------------------------------------|-----------|------------------------------|-----------------------------|------|
| <i>Factores de riesgo y síntomas</i> | pretest | 8 (6-11) | 8 (6-11) | 0,60 |
| | post-test | 9 (6-10) | 8 (6-9,5) | 0,36 |
| <i>Diagnóstico</i> | pretest | 2 (1-2) | 2 (1-2) | 0,83 |
| | post-test | 2 (1-2) | 2 (1-2) | 0,15 |
| <i>Monitorización</i> | pretest | 1 (1-2) | 1 (1-2,5) | 0,97 |
| | post-test | 2 (1-3) [#] | 2 (1-2) | 0,24 |
| <i>Tratamiento</i> | pretest | 2 (1-4) | 3 (2-4) | 0,09 |
| | post-test | 4 (3-5) [#] | 4 (3,5-5) [#] | 0,48 |

Los datos están expresados en medianas y percentiles 25-75. Test de la U de Mann-Whitney (diferencias entre grupos). (#) Test de Wilcoxon para 2 muestras apareadas, $P < 0,05$ intragrupo.

No se encontraron diferencias estadísticamente significativas entre los grupos en el *pretest* ni en el *post-test* para ninguno de los campos. La comparación intra grupo de las puntuaciones del *pretest* y *post-test* mostró incrementos significativos en el grupo *clase magistral* en el campo 3 “*Monitorización*” (test de Wilcoxon, $P = 0,03$) y campo 4 “*Tratamiento*” ($P = 0,001$). El grupo *caso/problema* solo obtuvo incrementos significativos en el campo 4 “*Tratamiento*” ($P = 0,003$).

El análisis de los ítems de reconocimiento y razonamiento del campo 1 “*Factores de riesgo y síntomas*” no mostró diferencias estadísticamente significativas entre los grupos en el *pretest* ni en el *post-test*. La comparación intra grupo de las puntuaciones del *pretest* y *post-test* tampoco mostró diferencias significativas en ninguno de los dos grupos (Tabla 25).

Tabla 25. Puntuaciones en el pretest y post-test para los ítems de reconocimiento y razonamiento del campo 1 “Factores de riesgo y síntomas”.

| Campo 1 | Test | Grupo clase magistral (n=27) | Grupo caso/problema (n= 25) | P |
|----------------|----------------|------------------------------|-----------------------------|------|
| Reconocimiento | pretest | 6 (5-6) | 6 (5-6) | 0,50 |
| | post-test | 5 (5-6) | 6 (5-6) | 0,82 |
| | P (intragrupo) | 0,21 | 1,00 | |
| Razonamiento | pretest | 2 (0-5) | 4 (0-5) | 0,55 |
| | post-test | 3 (1-5) | 2 (0-4) | 0,28 |
| | P (intragrupo) | 0,82 | 0,26 | |

Los datos están expresados en medianas y percentiles 25-75. Test de la U de Mann-Whitney (diferencias entre grupos). Test de Wilcoxon para 2 muestras apareadas (diferencias intragrupo).

No hubieron diferencias significativas entre los dos grupos al comparar la diferencia en las puntuaciones “post-test menos pretest” en ninguno de los campos (Tabla 26A) ni para los *ítems* de reconocimiento y razonamiento del campo 1 “Factores de riesgo y síntomas” (Tabla 26B).

Tabla 26A. Incrementos absolutos de puntuación para cada campo del conocimiento medido.

| Post-test menos pre-test | Grupo clase/magistral (n = 27) | Grupo caso/problema (n = 25) | P |
|-----------------------------|--------------------------------|------------------------------|------|
| Campo 1 post – Campo 1 pre | 1 [(-2) - 2] | -1 [(-2)-1,5] | 0,51 |
| Campo 2 post – Campo 2 post | 0 [(-1)-1] | 0 (0-1) | 0,45 |
| Campo 3 post – Campo 3 pre | 0 (0-1) | 0 [(-1)-1] | 0,23 |
| Campo 4 post – Campo 4 post | 2 (0-2) | 1 (0-2) | 0,39 |

Tabla 26B. Incrementos absolutos de puntuación del campo 1 diferenciados en los ítems de reconocimiento y razonamiento.

| Campo 1 post – campo 1 pre | Grupo clase/magistral (n = 27) | Grupo caso/problema (n = 25) | P |
|--|--------------------------------|------------------------------|------|
| Reconocimiento post – reconocimiento pre | 0 [(-1)-0] | 0 [(-1)-1] | 0,32 |
| Razonamiento post – razonamiento pre | 1 [(-1)-2] | 0 [(-2)-0,5] | 0,28 |

Los datos están expresados en medianas y percentiles 25-75. Test de la U de Mann-Whitney (diferencias entre grupos).

No encontramos diferencias entre los grupos en el número de participantes que mejoraron sus puntuaciones después de la enseñanza (Tabla 27). Los porcentajes de participantes que mejoraron sus puntuaciones en el campo 4 “*Tratamiento*” fueron relevantes (> 50%) en ambos grupos. Encontramos diferencias significativas en el *grupo clase magistral* al comparar el porcentaje de participantes que mejoraron sus puntuaciones entre los distintos campos de conocimiento medidos ($P = 0,005$).

Tabla 27. Número de participantes que obtuvieron puntuaciones mejores (rangos positivos), peores (rangos negativos) o iguales (rangos iguales) en las evaluaciones.

| Diferencias de puntuación | Rangos | Grupo clase/magistral (n = 27) | Grupo caso/problema (n = 25) | P |
|----------------------------|----------|--------------------------------|------------------------------|------|
| Campo 1 post – Campo 1 pre | Positivo | 14 (51,9) | 10 (40) | 0,39 |
| | Negativo | 11 (40,7) | 13 (52) | 0,41 |
| | Igual | 2 (7,4) | 2 (8) | 1,00 |
| Campo 2 post – Campo 2 pre | Positivo | 7 (25,9) | 7 (28) | 0,86 |
| | Negativo | 7 (25,9) | 4 (16) | 0,38 |
| | Igual | 13 (48,1) | 14 (56) | 0,57 |
| Campo 3 post – Campo 3 pre | Positivo | 9 (33,3) | 9 (36) | 0,84 |
| | Negativo | 3 (11,1) | 9 (36) | 0,03 |
| | Igual | 15 (55,6) | 7 (28) | 0,04 |
| Campo 4 post – Campo 4 pre | Positivo | 18 (66,7) [#] | 16 (64) | 0,84 |
| | Negativo | 4 (14,8) | 4 (16) | 1,00 |
| | Igual | 5 (18,5) | 5 (20) | 1,00 |

Valores expresados como número de casos (porcentaje intragrupo). (P) Test de *Chi-cuadrado* (diferencias entre grupos). (#) Test de Q de Cochran (diferencias intragrupo para rangos positivos).

El porcentaje de participantes que mejoró su puntuación de los *ítems* de razonamiento después de la enseñanza fue significativamente mayor en el *grupo clase magistral* que en el *grupo caso/problema* ($P = 0,03$). El porcentaje de participantes que mejoró su puntuación en los *ítems* de razonamiento fue relevante (> 50%) en el *grupo clase magistral*. En el *grupo clase magistral*, encontramos diferencias significativas entre el porcentaje de participantes que mejoró su puntuación en el reconocimiento de los *ítems* del campo 1 “*Factores de riesgo y síntomas*” y el porcentaje que mejoró la puntuación en el razonamiento de estos *ítems* ($P = 0,004$), (Tabla 28).

Tabla 28. Número de participantes que obtuvieron puntuaciones mejores (rangos positivos), peores (rangos negativos) o iguales (rangos iguales) en las evaluaciones del campo 1 "Factores de riesgo y síntomas".

| Diferencias de puntuación Campo 1 | Rangos | Grupo clase/magistral (n = 27) | Grupo caso/problema (n = 25) | P |
|--|----------|--------------------------------|------------------------------|------|
| Reconocimiento post – reconocimiento pre | Positivo | 5 (18,5) | 7 (28) | 0,41 |
| | Negativo | 11 (40,7) | 7 (28) | 0,33 |
| | Igual | 11 (40,7) | 11(44) | 0,81 |
| Razonamiento post – razonamiento pre | Positivo | 14 (51,9) [#] | 6 (24) | 0,03 |
| | Negativo | 10 (37) | 12 (48) | 0,42 |
| | Igual | 3 (11,1) | 7 (28) | 0,12 |

Valores expresados como número de casos (porcentaje intragrupo). (P) Test de *Chi-cuadrado* (diferencias entre grupos). (#) Test de McNemar (diferencias intragrupo para rangos positivos).

Encontramos correlación lineal entre el reconocimiento y el razonamiento de los ítems del campo 1 "Factores de riesgo y síntomas" antes de la enseñanza en el grupo caso/problema (Figura 22). El grado de la correlación fue leve ($r^2 = 0,261$).

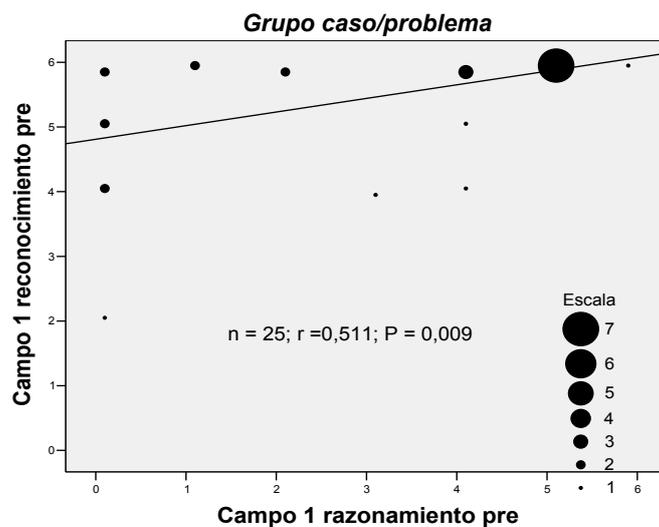


Diagrama de dispersión. Los puntos representan número de casos según la escala adjunta.

Figura 22. Correlación entre el reconocimiento y el razonamiento de los ítems del campo 1 "Factores de riesgo y síntomas" en el grupo caso/problema. (r) = coeficiente de correlación de Spearman.

6. DISCUSIÓN

Los tres estudios realizados no confirmaron la hipótesis inicial, el nuevo método docente basado en la *discusión de casos/problema* fue igual de eficaz pero no superior al método tradicional basado en la *clase magistral* para las habilidades cognitivas e integrativas evaluadas. Nuestros hallazgos fueron similares para los tres niveles de formación estudiados: *pre-grado*, *post-grado* y *Formación Médica Continuada*. No hemos encontrado estudios similares en Anestesiología que nos permitan comparar nuestros resultados. Estudios controlados y randomizados en otros campos de la educación médica muestran resultados concordantes con los de nuestros trabajos [76,131]. Creemos, por lo tanto, que para temas muy bien establecidos en Anestesiología en los tres niveles de formación, tanto la *clase magistral* como la *discusión de casos problema* son alternativas válidas para la cumplimentación de los objetivos docentes fijados. No podemos recomendar el empleo sistemático de un método docente sobre el otro. El diseño aplicado da fortaleza a nuestros resultados en los contextos y temas analizados.

El porcentaje de participación de los alumnos en los tres estudios fue similar a los porcentajes publicados en otros estudios semejantes [76,149,150], La elevada participación sugiere que los resultados fueron representativos de las poblaciones de alumnos analizadas; el diseño controlado de los estudios favoreció, probablemente, este hecho. Creemos que la participación fue menor en el estudio sobre la *valoración pre-anestésica* por la dificultad de acceso a la sesión docente de los residentes de primer año de Anestesiología de fuera de la ciudad.

Son muchos los estudios que han comparado el aprendizaje basado en problemas con métodos más tradicionales de enseñanza. Sin embargo, la interpretación de los resultados tiene que ser prudente. En la literatura no existe evidencia consistente de que el aprendizaje basado en problemas sea superior a la *clase magistral* en mejorar el conocimiento de los alumnos [49,50]. Dado que los recursos en educación son limitados, algunos autores opinan que la adopción de estos programas sólo estaría justificada si hay evidencia de que mejoran las competencias de los profesionales [92,151]. En dos revisiones recientes [16,152], el aprendizaje basado en problemas comparado con la enseñanza tradicional demuestra, con evidencia de moderada a fuerte, mejorar las competencias psicosociales de los profesionales que recibieron aprendizaje basado en problemas en formación médica de pre-grado como es enfrentarse a las dudas, habilidades de comunicación, comprensión de los contextos sociales de los pacientes o conciencia de las propias limitaciones sin diferencias en las competencias globales, técnicas, investigadoras, de dirección, de enseñanza, ni en los conocimientos adquiridos. Uno de los problemas principales es que muchos estudios no suelen incluir técnicas de evaluación que midan competencias más allá de la mera adquisición de conocimientos y que por lo tanto no captan las ventajas potenciales del aprendizaje basado en problemas. Como afirma Baños en una editorial reciente [15] el resultado de los estudios dependerá de lo que se esté evaluando. Como métodos de evaluación válidos para medir la eficacia del aprendizaje basado en problemas se han sugerido la auto-reflexión, la observación por expertos, la evaluación clínica objetivo estructurada (*ECOÉ*), los portafolios y las evaluaciones basadas en la resolución de casos/problema [96]. Otras dificultades incluyen que el grupo control no siempre es el mismo y muchos estudios son de poca calidad [153].

Otro aspecto importante a considerar es la necesidad de definir con claridad el término aprendizaje basado en problemas. Como hemos comentado previamente el aprendizaje basado en problemas y el *aprendizaje basado en la discusión de casos/problema* no son lo mismo. Aunque ambos son enfoques de aprendizaje centrados en el alumno, su diferencia más importante está en el papel del profesor, que en el *aprendizaje basado en la discusión de casos/problema* dirige y retroalimenta las discusiones [62]. Gracias a este enfoque del aprendizaje más dependiente del profesor, el método *aprendizaje basado en la discusión de casos/problema* es menos susceptible de la disfunción del grupo, un problema bien documentado e inherente al aprendizaje basado en problemas [62]. A diferencia del aprendizaje basado en problemas, *el aprendizaje basado en la discusión de casos/problema* no requiere muchos cambios organizativos y los costes económicos y recursos de personal que precisa son menores [21,36,49,62,154].

La mayoría de los estudios publicados sobre el *aprendizaje basado en la discusión de casos/problema* son unidimensionales, donde la investigación se centra en el *aprendizaje basado en la discusión de casos problema* como único método docente [60,68,77,155]. Son pocos los estudios en la literatura que comparen *el aprendizaje basado en la discusión de casos problema* y la *clase magistral*; la mayoría de ellos, además, sólo analizan la opinión de los alumnos o su cambio de actitud ante la enseñanza, niveles de resultados 1 (reacción) y 2A (aprendizaje, cambio de actitudes) del modelo modificado de Kirkpatrick para la evaluación de los resultados de la enseñanza [95], y están limitados por los tamaños pequeños de las muestras de alumnos [156-158]. Estas publicaciones confirman que el *aprendizaje basado en la discusión de casos/problema* es valorado positivamente por los alumnos, lo consideran un método más ameno, que aproxima mejor el conocimiento a la práctica clínica y perciben que sirve para mejorarles el razonamiento clínico, la interpretación de diagnósticos y la capacidad de pensar de forma más lógica. Los alumnos consideran que les ayuda a ser más participativos y los motiva más y los profesores piensan que es un método que mejora la capacidad de los alumnos de resolver problemas. Sin embargo, las *clases magistrales* son consideradas por los alumnos de más ayuda para prepararse los exámenes [156] y por los profesores como un método que fija más la atención del alumno y permite cumplimentar mejor los objetivos docentes [158].

Solo hemos encontrado en la literatura dos estudios controlados y randomizados [45,76] y un estudio comparativo [130] que incluyan otras medidas de nivel superior al 1 y 2A del modelo modificado de Kirkpatrick en el análisis de los resultados. En un estudio reciente, Wahlgren et al. [76] comparan la enseñanza de dermatología y venerología a los estudiantes de medicina con el método convencional y con el *aprendizaje basado en la discusión de casos/problema* simulados por ordenador mediante un sistema interactivo donde el alumno recibe la retroalimentación mediante la comparación con un especialista. Los autores demuestran, de acuerdo a otras publicaciones [159-161], que las nuevas tecnologías son compatibles con el método *aprendizaje basado en la discusión de casos/problema*. Aunque los alumnos valoran positivamente el método obtienen los mismos resultados de conocimientos en los exámenes de respuesta múltiple o abierta. El método convencional incluía *clases magistrales*, seminarios, talleres de trabajo y enseñanza a pie de la cama del paciente; los autores no diferenciaron entre ellos lo cual pudo influir en los resultados.

En el segundo estudio Hay y Katsikitis [45] demuestran que las puntuaciones en conocimientos sobre el tema de trastornos alimentarios de los estudiantes de medicina mediante exámenes de respuesta corta son mayores cuando el profesor que dirige la *discusión de los casos* es un docente experto en el método y contenido impartido comparado con el no experto. Las conclusiones refuerzan el papel que tiene el profesor en esta metodología. Por último, Mikail et al. [130] compararon la enseñanza de las implicaciones de las hierbas medicinales mediante una clase de *discusión de casos/problema* o con sólo un texto escrito. Las puntuaciones en el test de conocimientos después de la enseñanza se incrementaron el doble en el *grupo caso/problema*. Sin embargo los resultados de este estudio tienen una validez limitada: un texto no es equiparable a una *clase magistral* como método de enseñanza, el criterio de selección estuvo sesgado ya que se incluyeron en el “grupo texto” los alumnos que no asistieron a la clase de discusión de los casos y el tamaño muestral fue insuficiente, con solo 8 alumnos en el “grupo texto”.

Mientras no exista evidencia clara de la superioridad de un método docente al otro, no está justificado, por lo tanto, la supresión de ninguno de ellos.

Nuestros resultados confirman, por una parte, que con un diseño adecuado, el método *aprendizaje basado en la discusión de casos/problema* puede cumplir los mismos objetivos docentes y los alumnos obtener las mismas puntuaciones en las *evaluaciones basadas en casos* que la *clase magistral*. Una de las fortalezas del *aprendizaje basado en la discusión de casos/problema* es que puede ser igual de eficaz con formatos más sencillos, como los que aplicamos en el estudio de *postgrado* y *formación continuada*, o formatos que incorporen más soportes tecnológicos, como el que aplicamos en el estudio de *pregrado*. [72]. Por otro lado, nuestros resultados justifican igualmente el empleo de la *clase magistral* como método que favorece no sólo la adquisición de conocimiento, sino también la integración del conocimiento básico y clínico, la capacidad de aplicar el conocimiento para la solución de problemas, el manejo de la información o el uso del razonamiento clínico. Estos resultados desmitifican la mala imagen de la *clase magistral* a favor de las nuevas metodologías docentes, muchas veces sobrevaloradas en base a opiniones más que a datos objetivos.

Las comparaciones entre la *clase magistral* y métodos de enseñanza más novedosos (modalidades de aprendizaje basado en problemas, simulaciones, enseñanza por ordenador, videos interactivos etc) debe tratarse con cautela; mientras estos métodos suelen planificarse rigurosamente y son evaluados de forma sistemática no ocurre igual con las *clases magistrales*. La mayoría de los estudios asumen que las *clases magistrales* son fenómenos estables. En la práctica existe una gran variedad de enfoques dentro de este método que pueden convertirlo en un método de enseñanza potencialmente eficaz o ineficaz [3]. Parker [162] justifica el uso de la *clase magistral* y afirma que quienes estigmatizan este método docente lo que realmente están haciendo es rechazar las malas *clases magistrales*.

Existen excelentes guías sobre como mejorar las *clases magistrales* [3]. Una buena preparación y comienzo, un contenido estructurado que refuerze los puntos clave y cumpla los objetivos docentes, una buen formato audiovisual, una explicación clara, interesante y persuasiva, estar atentos a la audiencia y responder de acuerdo a ello, variar la actividad de los alumnos y sintetizar y resumir

favorecen la atención y el aprendizaje. La calidad del profesor es otro componente primordial para mejorar los resultados de la enseñanza [129]. En Anestesiología, en particular se han publicado varios artículos al respecto [129,163,164]. El contenido de la *clase magistral* aumenta su eficacia si se utiliza para estimular el debate e identificar modelos de auto-análisis y pensamiento crítico [129]; los temas relacionados con la morbi-mortalidad o los errores médicos, por ejemplo, tratados sin reparos mejoran la productividad de la clase [165].

El horario de las clases también juega un papel significativo: los residentes de Anestesiología que asisten a *clases magistrales* por la mañana obtienen mejores puntuaciones en los exámenes [7], sin embargo, los autores no encontraron correlación con el horario de tarde, el número de clases a la semana o la obligatoriedad de la asistencia.

La experiencia docente y clínica del profesorado que impartió las *clases magistrales* en nuestro estudio fue un factor que, sin duda, influyó en los resultados.

Respecto al contenido, el *soporte vital básico* y el *embolismo aéreo* intraoperatorio están íntimamente ligados con la morbi-mortalidad mientras que la *visita pre-anestésica* se relaciona de forma indirecta. El horario de las clases fue matinal para el *soporte vital básico (pregrado)* y a primera hora de la tarde para la *valoración pre-anestésica (post-grado)* y el *embolismo aéreo (formación continuada)*. Desconocemos si el factor horario afecta por igual al *aprendizaje basado en la discusión de casos/problema*; el hecho de que ambos métodos se impartieran al mismo tiempo minimizaría este efecto.

Un problema recurrente en la literatura es la falta de definición del método tradicional de enseñanza que puede englobar diversidad de formatos. En nuestro caso, las *clases magistrales* estaban bien estructuradas y su estilo no fue el de meras presentaciones orales, todas utilizaron el soporte técnico de una presentación Power Point, incorporaban información visual (ilustraciones, figuras etc) y algún diagrama y/o algoritmo. Para la enseñanza del *soporte vital básico* también se incluyó una grabación de video con ejemplos de actuación práctica. Todos estos aspectos han demostrado que favorecen el interés de los alumnos y potencialmente mejoran el aprendizaje [3] La calidad de las *clases magistrales* es un factor determinante de su eficacia. Creemos que definir con claridad los métodos docentes aplicados fortalece nuestros resultados.

Nuestros resultados no pueden extrapolarse a otras modalidades de aprendizaje basado en problemas ni a otros formatos de *aprendizaje basado en la discusión de casos/problema*. La variabilidad posible es alta; en nuestro caso tuvimos que adaptar el diseño a las condiciones del programa docente y del propio diseño del estudio. No podemos descartar que el método de *aprendizaje basado en la discusión de casos/problema* pudiera obtener mejores resultados aplicado a otros temas, en contextos diferentes o si hubiéramos analizado otros parámetros. En este sentido algunas publicaciones muestran la eficacia de los escenarios o casos/problema en mejorar las habilidades no técnicas de los alumnos (administrar tareas, trabajar en equipo, tomar decisiones etc) y el juicio crítico [77,116,149]. Las sesiones interactivas, a diferencia de la *clase magistral*, pueden incluso originar cambios moderadamente importantes en la práctica profesional de los alumnos [166,167]. El problema es que estas habilidades son mucho más difíciles de evaluar con los métodos

de evaluación tradicionales [168] y tampoco permiten inferir si el resultado final para el paciente será mejor [169]. Algunos autores han referido también que la retroalimentación que hace el profesor durante el *aprendizaje basado en la discusión de casos/problemas* favorece una mejor capacidad de razonamiento entre los alumnos [21,170]. Nuestros resultados no confirman esta conclusión.

El principio fundamental de la *educación basada en los resultados* [92] es saber si los métodos de enseñanza permiten a los alumnos alcanzar los objetivos docentes de los programas de formación y usar métodos de evaluación que lo demuestren. El aspecto importante de nuestro estudio subyace en el hecho de que con ambos métodos, uno más pasivo y otro más participativo, se pueden conseguir resultados igual de buenos y por lo tanto ambos métodos docentes constituyen opciones válidas. En nuestro caso, la elección podría supeditarse a las preferencias del profesor o de los propios alumnos [21,171]. Otra opción sería incluir los dos métodos en los programas de formación como métodos complementarios.

La universidad de McMaster ha introducido, a lo largo de sus 30 años de historia, cambios curriculares que han supuesto un mayor número de sesiones didácticas y un cambio de la metodología docente más próxima al *aprendizaje basado en la discusión de casos/problema* [12]. Estos cambios los han basado en la evidencia de sus resultados, en las necesidades sociales y en las necesidades de sus propios alumnos. En nuestra opinión, este hecho manifiesta claramente el interés renovado que despiertan ambos métodos de enseñanza y los beneficios potenciales de los programas híbridos. Discernir entre los dos métodos exigiría medir otras variables, habilidades o competencias. Haidet et al. [172] al comparar el aprendizaje pasivo (*lección magistral*) y el activo (sesión única de *discusión basada en problemas* dirigida por el profesor) no encuentran diferencias entre los dos métodos en la ganancia de conocimientos ni en el cambio de actitud de los residentes; sin embargo, a pesar de que el método de aprendizaje activo lo consideraron más atractivo, valoraron mejor la *lección magistral* en términos de cumplimentación de los objetivos docentes.

El balance coste/beneficio de los métodos de enseñanza, como los utilizados por nosotros en este estudio, es difícil de valorar porque depende de una organización pre-establecida y de la infraestructura disponible en los diferentes centros. En nuestro caso, ambos métodos fueron igual de asequibles dentro de nuestra organización. Disponer de diferentes alternativas docentes y saber adaptarlas a cada caso en particular es, probablemente, la mejor manera de enseñar y aprender las distintas competencias en Anestesiología. Creemos que es necesaria la diversidad de los métodos de enseñanza para la educación en medicina en general y en Anestesiología en particular.

El aprendizaje basado en problemas [33,173,174] y el *aprendizaje basado en la discusión de casos/problema* [53,55] se están introduciendo en los programas de formación de Anestesiología. Ambos métodos han demostrado ser aplicables y con un alto grado de satisfacción de los alumnos [53,54,175]. En un estudio reciente, comparando ambos métodos, los alumnos y profesores prefirieron mayoritariamente el *aprendizaje basado en la discusión de casos/problema* porque percibían un mejor aprovechamiento del tiempo: centraba más el conocimiento en el tema, la carga de trabajo era menor y ofrecía mayores oportunidades para la aplicación clínica de los conocimientos y habilidades adquiridas [21]. La satisfacción de los participantes con un determinado método de

enseñanza podría influir en su eficacia [176]. Nosotros no evaluamos este aspecto y esto constituye una limitación del estudio. Sin embargo, las encuestas de opinión tienen un valor limitado y están sujetas a sesgo y error. Se ha observado una correlación baja entre la autoevaluación y la evaluación objetiva de las competencias tanto en profesionales [177,178] como en estudiantes [179,180]; por este motivo, la satisfacción de los alumnos no puede ser extrapolada como predictor de la competencia profesional [49,181]. En un estudio reciente en Anestesiología [55] se demuestra que la percepción de los alumnos sobre la potencial mejoría de sus conocimientos y habilidades clínicas con un determinado método docente (simulación de alto realismo) tampoco coincide con los resultados obtenidos.

El alto grado de acuerdo entre evaluadores que encontramos en nuestro estudio es una medida de fiabilidad y capacidad de repetición de las evaluaciones aplicadas, que fortalece el diseño de las evaluaciones y podría favorecer la difusión de su uso. Hasta ahora no se ha descrito un método válido y universal para la evaluación de la eficacia de la enseñanza en *pregrado* [135,136], *postgrado* [88,99,104] o *Formación Médica Continuada* [182]. El objetivo final sería poder demostrar si el método docente mejora la práctica clínica. Por este motivo, es importante aproximar al máximo la evaluación a la práctica clínica. Se ha de definir con claridad el aspecto del aprendizaje o competencia que estamos midiendo. En Anestesiología, en particular, no hay referencias sobre cuál es el mejor método de evaluar los conocimientos [183]. Los exámenes de respuesta múltiple son el método de evaluación más comúnmente empleado, sobre todo por su medición objetiva de las respuestas. Sin embargo, no valoran la construcción de la respuesta por parte del alumno, su capacidad de aplicar el conocimiento, la resolución de problemas ni la competencia profesional [99,104]. A diferencia de los test de respuesta múltiple, el diseño de evaluaciones de respuesta abierta nos permite comprender mejor la construcción de la respuesta del alumno y el análisis de habilidades cognitivas más allá del simple reconocimiento [96]. La *evaluación basada en casos/problema* tal y como la hemos utilizado en la presente investigación es una herramienta válida para medir el razonamiento y la aplicación de los conocimientos para la solución de los problemas [88,96]. En 1995, la universidad de McMaster introdujo un examen de respuesta corta abierta en su programa curricular. Este cambio se fundamentó en entender que el manejo de los problemas clínicos comunes incluía áreas de conocimiento tales como el razonamiento clínico que eran esenciales para la competencia profesional del alumno [12].

Para medir la capacidad de aplicar en la práctica los conocimientos teóricos nos tenemos de valer de otro tipo de herramientas de evaluación [88,96,99,122]. Las dos más frecuentemente usadas en Anestesiología son la observación por expertos y la simulación. La observación de la actuación profesional por expertos está limitado por su fiabilidad; también se ha cuestionado si la observación de las decisiones de los alumnos muestran realmente el uso apropiado del conocimiento y el razonamiento [107]. Por su parte, los simuladores no están al alcance de todos los centros y universidades. Ambas evaluaciones están, además, condicionadas logísticamente por el tiempo y recursos que precisan [88,99]. Nosotros adoptamos un método de evaluación que se adaptara a los programas de formación, a los recursos disponibles y sobre todo, que no implicara una disminución del tiempo de enseñanza. El formato de caso clínico aplicado aproximó las evaluaciones a la realidad

clínica, y la clasificación de las preguntas por campos de conocimiento favoreció el seguimiento de los objetivos docentes.

El método de evaluación por campos de conocimiento, tal y como lo hemos aplicado en este estudio, no se ha utilizado previamente en Anestesiología. Sin embargo, el sistema de *evaluación basado en casos* ha demostrado ser válido para identificar factores de riesgo, mecanismos fisiopatológicos, seleccionar diagnósticos y reconocer tratamientos en otras áreas; su formato en forma de respuesta abierta se ha validado también para medir el razonamiento [96]. Además las preguntas agrupadas por *ítems*, de forma semejante a como lo hemos hecho nosotros, estimula habilidades cognitivas más complejas como la aplicación del conocimiento en la resolución de problemas [184]. Su formato estructurado y la delimitación con claridad de las áreas de las competencias que se quieren analizar son factores que han demostrado aumentar la fiabilidad de este método de evaluación [88]. Existen otros métodos de evaluación validados que examinan áreas de conocimiento práctico, como establecer un diagnóstico o un tratamiento, al modo de alguno de los campos de conocimiento de nuestras evaluaciones [185]. Nosotros medimos habilidades cognitivas (conocimiento, aplicación práctica del conocimiento) e integrativas (usar el razonamiento clínico o unir conocimiento básico y clínico) conforme a la clasificación de las competencias profesionales de Epstein y Hundert [104]. Otros autores también han hecho sus propias adaptaciones modificadas para la evaluación de las competencias [16].

Las altas puntuaciones previas en ambos grupos en los temas “*valoración pre-anestésica*” y “*embolismo aéreo*” sugieren que los contenidos ya eran conocidos por la mayoría de los alumnos y, por lo tanto, los objetivos docentes en los niveles de formación de *post-grado* y *formación continuada* no deberían buscar incrementar los conocimientos absolutos de los alumnos, sino estimularlos a que sepan aplicar este conocimiento en situaciones prácticas. Sería interesante adaptar los objetivos docentes a las necesidades de los alumnos en base a los resultados de las evaluaciones antes de la enseñanza [186]. Estas altas puntuaciones limitaron además la posibilidad de detectar mejorías significativas en las evaluaciones. El análisis de otras habilidades o competencias hubiera podido aportar la ventaja de detectar mejor las diferencias entre los dos métodos de enseñanza.

Un factor que pudo favorecer los buenos resultados del método *aprendizaje basado en la discusión de casos/problema* fue el entorno realístico de la enseñanza. Aproximar la enseñanza a la realidad favorece el aprendizaje [72,187]: los alumnos se implican más e interpretan mejor sus experiencias y conocimientos. Este hecho podría explicar porque cuando se compara el aprendizaje basado en problemas con la simulación no se encuentran diferencias significativas en el nivel de conocimientos y habilidades prácticas de los alumnos [55]. En nuestro estudio sobre *soporte vital básico*, hemos demostrado que también es posible incorporar soportes tecnológicos al método tradicional de *clase magistral* que aproximen la enseñanza a la realidad y puedan así mejorar su eficacia. Reiteramos nuestra opinión de que los estudios de investigación docente han de describir con claridad las características de los métodos de enseñanza para poder interpretar correctamente los resultados.

Es interesante remarcar que la mayoría de alumnos no alcanzaron las puntuaciones máximas en las evaluaciones post-enseñanza. Estos resultados reflejan la dificultad de alcanzar los objetivos

docentes en una sola estación de trabajo clase o sesión. Este dato ya ha sido referenciado en la literatura [188,189] y justifica la recomendación de la *Formación Médica Continuada* [138,188,190].

No descartamos la posibilidad de que existan otros métodos de enseñanza más eficaces que la *clase magistral* o el *aprendizaje basado en la discusión de casos/problema* para los temas analizados. Los programas de entrenamiento basados en simulaciones de alta fidelidad en anestesia han demostrado ser efectivos en mejorar las habilidades no técnicas como la gestión del trabajo, el trabajo en equipo o la capacidad de decisión [116]; el entorno realístico del aprendizaje parece ser un factor determinante [191]. Los participantes también perciben un cambio en su práctica clínica como consecuencia de este tipo de formación [192,193], aunque esta percepción no implica mejores resultados en la adquisición de conocimientos teóricos medidos por examen de respuesta múltiple y en las habilidades clínicas evaluadas en un simulador cuando se compara con la *discusión de casos problema* [55]. Otros autores sí han referido mejor adquisición y retención del conocimiento y mejores habilidades prácticas comparando la simulación con las *clases magistrales* [194]. Estos programas conllevan un elevado coste económico, difícil de obtener, porque requieren un entorno físico sofisticado, equipamiento y personal entrenado [195,196]. Aunque la evidencia indica que los métodos de simulación son efectivos para mejorar las habilidades técnicas y psicomotoras [197], en el momento actual faltan estudios de calidad y rigor suficiente que justifique la implementación de la enseñanza basada en simuladores en educación médica en general [198-200] y en Anestesiología en particular [201,202]; probablemente sea necesario, como ha ocurrido con el aprendizaje basado en problemas, desarrollar formatos más sencillos, funcionales y asequibles y analizar cuál es su aplicabilidad y eficacia en áreas concretas de la anestesia [196].

Tampoco comparamos el *aprendizaje basado en la discusión de casos/problemas* con el aprendizaje basado en problemas u otras variantes de este método. La enseñanza con casos cortos esquematizados ha demostrado mejorar la resolución de problemas rutinarios en la práctica clínica comparado con el aprendizaje basado en problemas [203]. El programa de formación médica de *pregrado* de la Universidad de McMaster ha cambiado igualmente el formato de casos largos y con múltiples objetivos por casos mucho más cortos y focalizados. Esto ha permitido al alumno abordar conceptos importantes a partir de distintas situaciones clínicas y ha ayudado al profesor a conducir mejor las sesiones y a mejorar la retroalimentación [12]. En opinión de algunos autores el *aprendizaje basado en la discusión de casos/problema* es la progresión natural del aprendizaje basado en problemas [68]. La estrecha franja de tiempo para la enseñanza dentro de cada programa de formación fue otro condicionante del diseño del estudio. Nosotros intentamos mejorar la eficacia de la enseñanza sin cambios organizativos sustanciales. En este sentido, los resultados de nuestro estudio justifican que se pueda instaurar el *aprendizaje basado en la discusión de casos/problema* en nuestros programas de formación.

Nuestro estudio presenta varias limitaciones. El estudio trató aspectos limitados del conocimiento (sólo habilidades cognitivas e integrativas) en temas muy concretos (*soporte vital básico, valoración pre-anestésica y embolismo aéreo*) y en un escenario limitado (una única estación de trabajo, sesión o clase de tiempo limitado con pocos alumnos). No podemos extrapolar, por tanto, nuestros

resultados y conocer cual hubiera sido el efecto sobre el aprendizaje si los métodos de enseñanza se hubieran aplicado durante todos los programas. Limitaciones institucionales, tanto en disponibilidad de tiempo como recursos, restringieron el diseño del estudio al que realizamos en el presente trabajo. Sería importante en futuros estudios valorar la generalización de nuestros resultados a otros temas. Limitar el análisis a un solo tema, específico del nivel de formación, nos ayudó, por otro lado, a controlar mejor las condiciones del estudio. Los estudios randomizados y controlados a nivel curricular están limitados por la imposibilidad de estandarizar las intervenciones docentes [204]. Los estudios de investigación docente que evalúan la competencia general, nivel “*hacer*” de la *pirámide de Miller* requieren un diseño de estudio y unas técnicas de evaluación totalmente diferentes a los planteados en esta investigación [107,205]. Dicho nivel de evaluación esta enfocado en evaluar la progresión del alumno a lo largo de su formación curricular y precisa, por tanto, de múltiples métodos de evaluación aplicados a lo largo del tiempo; desafortunadamente esta metodología tampoco está exenta de limitaciones [122].

Las puntuaciones de nuestras evaluaciones no pueden predecir la actuación profesional o el éxito de los alumnos en el *soporte vital básico*, la *visita pre-anestésica* o el manejo del *embolismo aéreo* intraoperatorio. Esta limitación ha sido referida en la literatura como una de las limitaciones principales de los estudios de investigación docente como el nuestro [56]. Sin embargo, Reich et al. [206] encontraron que una mala práctica de los residentes de anestesia sí se asociaba con factores de personalidad (introversión marcada, poca flexibilidad) y déficits en determinadas habilidades cognitivas y académicas (dificultad en realizar tests aritméticos de respuesta rápida que requerían dividir la atención, errores en los tests de detección visual, puntuaciones bajas en el test de conocimientos iniciales).

Nuestro método de *evaluación basado en casos* no permitió medir el resultado de la enseñanza más allá del nivel 2B de aprendizaje (modificación de los conocimientos o habilidades) del modelo modificado de Kirkpatrick [95] ni nos permitió medir las competencias más allá del nivel “*saber cómo*” de la *pirámide de Miller* [106]. Creemos que es importante que los estudios de investigación como el nuestro describan con claridad cuales son los resultados de la enseñanza, y por tanto las competencias, que estan midiendo con los métodos de evaluación aplicados. Un diseño que incluya los efectos de los métodos de enseñanza en más habilidades, técnicas, contextuales, de relación etc, permitiría una valoración más aproximada de la competencia de los alumnos en la práctica clínica [107,119,168,207]. La ausencia de diferencias en las habilidades cognitivas e integrativas con los dos métodos de enseñanza aplicados no excluye, por lo tanto, la posibilidad de que existan diferencias en las habilidades prácticas. Nosotros desconocemos si la enseñanza ayudó a mejorar la competencia clínica de los alumnos, nivel máximo de resultados de la enseñanza de Kirkpatrick y de competencia de la *pirámide de Miller*, para poder afirmarlo deberíamos disponer de instrumentos de evaluación de la actuación médica en la situación clínica real [208]. Estos métodos de evaluación son difíciles de aplicar [111]; los simuladores son una alternativa [181] pero también limitan el análisis de la competencia a un nivel inferior (“*mostrar como*” de la *pirámide de Miller*) [88].

La eficacia de los distintos métodos docentes esta influenciada por el esfuerzo que generan en los alumnos por conseguir superar las pruebas de evaluación [89]. El hecho de que nuestros estudios no tuviesen influencia en la puntuación académica y la participación fuese voluntaria y anónima redujo el potencial sesgo de las respuestas de los alumnos. Una de las ventajas del *aprendizaje basado en la discusión de casos/problema*, y que lo diferencia del aprendizaje basado en problemas, es que la enseñanza puede realizarse en el entorno de una sola sesión, igual que la *clase magistral*, lo que permite rentabilizar mejor el tiempo de dedicación del alumno y el profesor sin mermar necesariamente la eficacia de la enseñanza [21,62].

El efecto de los métodos docentes sobre la retención del conocimiento en el tiempo después de la enseñanza es otro aspecto crítico cuando se valora la eficacia [16]. Los estudios de seguimiento de la retención del conocimiento presentan dificultades logísticas y la pérdida de casos es una limitación frecuente en los estudios a medio y largo plazo [54,209]. En el estudio de formación continuada, nosotros fuimos incapaces de reclutar suficientes participantes 8 meses después para evaluar la retención de los conocimientos a pesar de ser la muestra de alumnos más controlada de los tres niveles de formación. Dos estudios recientes, uno en Anestesiología [210] y otro en enfermería [211], parecen indicar que el aprendizaje basado en problemas favorece la retención de conocimientos, sin embargo, estos resultados están limitados por el pequeño tamaño de las muestras. La revisión de la literatura no muestra evidencias convincentes de que las modalidades de aprendizaje basado en problemas sean más eficaces que la *clase magistral* en la retención de conocimientos [20,212].

Es posible que hubieran diferencias entre los dos métodos de enseñanza que no pudimos detectar porque fijamos una diferencia en las puntuaciones de al menos el 25 ó 30%. Estas diferencias son frecuentes en los estudios comparativos con metodologías de aprendizaje basado en problemas [55], aunque quizás fueron excesivas [78]. Para detectar diferencias menores hubiéramos tenido que incluir muchos más alumnos. A diferencia de los estudios clínicos donde es posible incluir más o menos pacientes, en los estudios docentes el tamaño ideal de la muestra está altamente restringido por los recursos y condicionantes existentes [213]; uno de los condicionantes es el número de alumnos adecuados al plan de formación. En nuestro caso, el tamaño de la muestra estuvo determinado por el número de alumnos que hicieron el curso en la Facultad de Medicina, el número de residentes de primer año de Cataluña que asistieron a la clase de *valoración pre-anestésica* o el número de anesestesiólogos inscritos en el curso de *Formación Médica Continuada*. Incluir alumnos de varios años es una opción; nosotros la aplicamos en el área de formación de *post-grado* porque consideramos que un máximo de dos años académicos consecutivos no incorporaban cambios curriculares en la formación de los residentes ni modificaciones en la metodología y experiencia del profesorado que pudieran sesgar los resultados; sin embargo, no siempre es posible, por ejemplo el curso de *formación continuada* incluido en este estudio sólo se repite cada 3 años. La alternativa de un estudio multicéntrico para incluir un mayor número de alumnos es compleja por la gran variabilidad existente entre los distintos centros de formación.

Un tamaño de muestra pequeño es uno de los impedimentos característicos de los estudios de educación médica, difícil de solucionar [56]. La mayoría de diseños de estudios randomizados no

incluyen información sobre cómo se calculó el tamaño muestral o no lo mencionan. [214]. Nosotros restringimos la randomización a dos bloques para que el número de casos en cada grupo fuese el máximo y optimizar así el poder estadístico de las mediciones. La hipótesis del estudio era unilateral (*caso/problema* mejor que *clase magistral*); sin embargo, para el cálculo del tamaño de la muestra consideramos Z_{α} para hipótesis bilateral, lo que resultó en un tamaño de muestra calculado mayor (Z_{α} bilateral = 1,96, Z_{α} unilateral = 1.645). A pesar de ello, los tamaños de muestra del presente estudio podrían ser considerados pequeños desde el punto de poder estadístico, sin embargo son relevantes en términos de representatividad especialmente si los comparamos con los datos publicados en la literatura [172,215].

El método de evaluación fue básico, con preguntas simples y de un rango de puntuación limitado. Esta simplicidad pudo causar un sesgo que afectase la validez y nos hace ser cautos a la hora de generalizar nuestros resultados. La simplicidad de las preguntas pudo haber condicionado el hecho de que no detectásemos diferencias entre los grupos. En nuestro caso pudo influir en la valoración de los campos de conocimiento que sólo incluían uno o dos *ítems*, pero creemos que no fué un factor determinante porque el mismo formato detectó puntuaciones bajas antes de la enseñanza y diferencias en las puntuaciones después de la enseñanza en ambos grupos. Nuestra herramienta de evaluación podría ser revisada para incluir un mayor número de *ítems* discriminatorios y de campos de puntuación, siempre y cuando el consecuente incremento de tiempo para la evaluación no mermara el tiempo de la enseñanza.

La enseñanza es difícil de valorar. Norman y Schmidt [216] son contundentes cuando afirman que sólo pueden medirse resultados sencillos de la enseñanza. En su opinión, la medición de resultados complejos está muy distorsionada por las múltiples variables que interactúan, lo que impide conocer el efecto real de los métodos de enseñanza. Se ha demostrado que la aplicación de algoritmos sencillos, explícitamente estructurados tanto para el *soporte vital básico* como para el *embolismo aéreo* durante la anestesia favorecen su aprendizaje, su aplicabilidad, que el diagnóstico sea más precoz y el manejo terapéutico mejor [139,217]. Otros estudios de investigación docente en Anestesiología han aplicado algoritmos similares a los nuestros [55]. Preguntas simples no implican respuestas simples; en nuestro diseño las respuestas de los alumnos reflejaron la capacidad de los alumnos de integrar y aplicar las habilidades cognitivas. Cada campo de conocimiento medido corresponde a un proceso diferente de aplicación e integración del conocimiento. No calculamos una puntuación global para las evaluaciones porque, en nuestra opinión, las puntuaciones de los diferentes campos medidos no eran comparables. La escasa correlación que encontramos entre la habilidad de reconocer y razonar después de la enseñanza refuerza esta idea. Diferenciar y agrupar los campos de conocimiento medidos en función de distintas habilidades obliga a la sobreponderar los campos con un número más bajo de *ítems* con el consiguiente riesgo de sesgo. El número bajo de *ítems* de la mayoría de los campos nos limitó la posibilidad de ampliar el estudio de correlación. Los *ítems* con un 100% de acierto o máxima puntuación después de la enseñanza en ambos grupos nos impidieron discriminar el efecto del método de enseñanza; desafortunadamente el bajo número de casos incluido en los estudios piloto de *post-grado* y *formación continuada* no nos permitió ponderar la dificultad de los *ítems*. Algunos autores [55] han aplicado sistemas de puntuación ponderada de

acuerdo a la importancia de cada *ítem* en opinión de varios expertos (*método Delphi*). Esta metodología se ha utilizado para el cálculo de puntuaciones globales. Nosotros puntuamos por igual las respuestas correctas de cada uno de los *ítems* pero compensamos las posibles diferencias relativas agrupando los *ítems* por campos de conocimiento que analizamos por separado. Con esta clasificación consideramos que los *ítems* de cada campo quedaron compensados y por lo tanto no precisaban que sus puntuaciones fuesen ponderadas.

La evaluación previa a cualquier método de enseñanza es esencial para valorar su eficacia y, si se lleva a cabo con suficiente antelación, también permite ajustar los contenidos docentes a las características del grupo y mejorar así los objetivos de aprendizaje. En nuestro estudio, las evaluaciones se realizaron inmediatamente antes y después de la enseñanza para asegurar que los alumnos no recibían ninguna información sobre el tema excepto la recibida en la propia estación de trabajo, clase o sesión docente. A pesar de que la adquisición inmediata de habilidades cognitivas e integrativas es sólo una pequeña parte del complejo proceso del aprendizaje profesional, nos ofrece la posibilidad de atribuir los resultados obtenidos al método de enseñanza aplicado sin contaminaciones externas. En una editorial reciente HS Barrows remarca la importancia de este diseño de evaluación para mejorar su utilidad [181].

Existe la posibilidad de que las evaluaciones previas pudieran influir en la mejoría de las puntuaciones después de la enseñanza debido al hecho de practicar los *ítems*, un efecto referenciado en las evaluaciones docentes [218]. En los estudios de *post-grado* y *formación continuada* mantuvimos el mismo formato de evaluación pero con casos distintos antes y después de la enseñanza. En el estudio de *pregrado* repetimos la misma evaluación antes y después de la enseñanza y no podemos descartar, por tanto, este posible sesgo; sin embargo, las puntuaciones después de la enseñanza no fueron máximas, lo que creemos reduce la posibilidad de este efecto. No tenemos conocimiento de otros estudios controlados y randomizados en la formación de la Anestesiología que incluyan evaluaciones de habilidades cognitivas e integrativas antes y después de la enseñanza. Por consiguiente, no podemos comparar nuestros resultados.

Una de las dificultades de los estudios de investigación como el nuestro es la homogeneidad entre los grupos analizados; nuestro criterio de selección de controles basado en un diseño controlado y randomizado en el mismo entorno y tiempo de la clase y la evaluación antes de la enseñanza limitan estas dificultades. La ausencia de diferencias entre los grupos en su formación o experiencia previa, los hospitales de procedencia o en las puntuaciones antes de la enseñanza son indicadores de homogeneidad de nuestra muestra de alumnos. En el estudio en formación *post-grado* encontramos heterogeneidad entre los grupos antes de la enseñanza para uno de los campos medidos; este hecho refuerza la importancia de que en educación médica es necesario realizar una evaluación previa para poder valorar objetivamente cualquier método docente. Nosotros sólo registramos la formación y experiencia previa de los alumnos de *pregrado*; en *post-grado* y en *formación continuada* no recogimos estos datos.

Nuestra decisión se basó en que la enseñanza del *soporte vital básico* tiene una amplia difusión social, más allá del entorno de la facultad de Medicina. Por su parte, la *valoración pre-anestésica* es

un tema específico que no se explica fuera de los programas de formación de los residentes de anestesia. Respecto al *embolismo aéreo* en anestesia, los casos planteados en las evaluaciones del curso de *formación continuada* son incidentes críticos trascendentes para la formación en Anestesiología pero de incidencia escasa; en nuestra opinión, sólo los anesthesiólogos que dedican su actividad asistencial regular a áreas de riesgo de aparición de esta complicación (verbigracia, anestesia en neurocirugía mayor) tendrían un nivel de experiencia capaz de influir en los resultados, y este no es el perfil de profesionales que se inscriben en los cursos del Comité para la Educación Europea en Anestesiología.

La formación previa en reanimación de los alumnos de medicina pudo condicionar los resultados obtenidos. En el diseño del estudio estimamos que un porcentaje no despreciable de alumnos habrían recibido algún tipo de formación previa en reanimación. Nuestros resultados (32% de los alumnos) confirmaron nuestra presunción. No excluimos del estudio estos alumnos porque nos hubiera dificultado alcanzar el tamaño de muestra calculado. Afortunadamente estos alumnos se distribuyeron por igual en ambos grupos; además, la comparación de las evaluaciones antes y después de la enseñanza evitó el potencial efecto de la formación previa sobre los resultados.

No se ha podido demostrar que el lugar de trabajo, los años de experiencia profesional o la universidad de origen de los alumnos condicionen el éxito del aprendizaje; sin embargo, el sexo femenino sí ha demostrado mejorar la adquisición de conocimientos teóricos y la actuación profesional en la práctica clínica [219] tanto con las *clases magistrales* como con el *aprendizaje basado en la discusión de casos/problema* [220]; en nuestro caso, la ausencia de diferencias entre los grupos evitó este posible sesgo.

No registramos la edad de los alumnos de *formación médica continuada*; desconocemos por tanto si fue similar en ambos grupos y si este dato pudo condicionar nuestros resultados. Sin embargo, la edad, a diferencia del sexo, no ha demostrado ser un factor predictivo del aprendizaje en *Formación Médica Continuada* [219].

La imposibilidad de diseñar estudios a doble ciego es una limitación innata a este tipo de estudios [204,216,221]. Consideramos muy difícil, por tanto, poder excluir el potencial efecto del profesor o el de los propios alumnos sobre los resultados al saberse observados [222].

En los tres programas de formación del estudio, los alumnos estaban acostumbrados al método tradicional de *clase magistral* mientras el *aprendizaje basado en la discusión de casos/problema* fue una novedad. Llama la atención que los alumnos que realizaron una actividad docente desconocida adquirieran habilidades similares a los que recibieron una *clase magistral*, ya que la mayoría hemos sido educados para tener éxito en métodos docentes tradicionales. Desconocemos si un mayor tiempo de adaptación en el *grupo caso/problema* hubiera mejorado los resultados como sugieren algunos autores [78].

Aunque el número reducido de alumnos en el *grupo clase magistral* pudo mejorar los resultados, no se ha podido demostrar que un grupo pequeño, como factor independiente, mejore los resultados de

la *clase magistral* [223]. Respecto al método *aprendizaje basado en la discusión de casos/problema* ya ha sido empleado con éxito en un tamaño de grupo moderado como el nuestro [58,60].

Que uno de los grupos mejorase o empeorase más los *items* de un campo sugiere que cada profesor pudo insistir más o menos en dichos aspectos durante su explicación. En estudios de investigación docente como el nuestro, interactúan multitud de variables difíciles de controlar: el contexto donde se realiza la enseñanza, el efecto ya comentado de los profesores sobre los resultados y la motivación o las expectativas personales de los participantes; todas ellas pudieron interferir en los resultados [21,95,221,224]. A diferencia de los estudios clínicos, en el campo de la educación médica la divergencia es la norma, lo que dificulta enormemente las comparaciones. No podemos descartar la posibilidad de que nuestros resultados cambiaran al adaptar nuestra metodología a otros contextos porque se ha demostrado que pequeñas modificaciones en los enfoques docentes originales pueden alterar su eficacia [83].

Un aspecto diferencial de los dos métodos de enseñanza utilizados fue la naturaleza de la participación de los alumnos, pasiva durante la *clase magistral* en contraste con una participación más activa en la *discusión de casos/problema*. Sin embargo, a la luz de nuestros resultados, no podemos descartar algunos factores que pudieron haber contribuido a la similitud entre ambos métodos. El elevado nivel de conocimientos previos de los alumnos de *post-grado* y *formación continuada* es un factor importante capaz de atenuar las posibles diferencias debidas al método docente; de igual modo, estas diferencias pudieron estar atenuadas por el efecto discusión del tiempo concedido para preguntas y respuestas del *grupo clase magistral*. Esta práctica hace que la *clase magistral* sea más interactiva, puede renovar la atención del alumno, generar interés, darle oportunidades de pensar y reforzar su comprensión [3]. Incluir tiempo para preguntas y respuestas no es infrecuente con este método de enseñanza [215]. Incluir una discusión facilitada puede mejorar la puntuación de las evaluaciones [59]. Nosotros no medimos el porcentaje de interacción de los alumnos; sin embargo, creemos que un tiempo tan corto en el grupo *clase magistral* sólo pudo tener una influencia mínima sobre las habilidades medidas. Creemos que hubiera sido interesante disponer de datos objetivos sobre el porcentaje de participación de cada alumno en el *grupo caso/problema*. Limitaciones logísticas nos impidieron registrarlo. Tärnvik [58] grabó en video la discusión de los casos y comprobó que existían una gran variación entre los estudiantes en el número de intervenciones así como en el tiempo total consumían hablando. En opinión del autor, esta variación reflejaba una predisposición individual que debería ser respetada. Aunque no estudió la correlación entre el tiempo utilizado y los resultados observó que el alumno que menos tiempo había utilizado (50 segundos) dió su nombre en la evaluación y fue el que obtuvo la mejor puntuación.

Las limitaciones que nosotros hemos encontrado son las mismas limitaciones que se han atribuido a los estudios randomizados y controlados en la literatura según el enfoque de la *educación basada en la evidencia* [89,225-227]. Comparada con la investigación clínica, la investigación en educación puede ser más compleja, los factores de confusión más aparentes, los contenidos menos explícitos y, como hemos comprobado, los estudios controlados más difíciles. Además, el resultado de la educación sobre el bienestar del paciente y la calidad de la salud es menos directo que con los

procedimientos médicos. Estas dificultades han planteado la pregunta de si se ha de continuar dando soporte y seguir participando en este enfoque educativo. En una editorial reciente Norman [85] pone en entredicho la validez de la búsqueda de evidencia mediante la revisión sistemática y meta-análisis de estudios randomizados y controlados porque la gran heterogeneidad existente entre los estudios y medidas de evaluación hace imposible evitar los sesgos en la interpretación de los resultados, lo que acaba engrosando el *fenómeno de no diferencias significativas* [228] al comparar la eficacia de los métodos docentes. Según Norman, es prácticamente imposible obtener un resultado no adulterado que pueda atribuirse solo a la intervención docente. Sin embargo, muchos de los problemas de la investigación docente también subyacen en la investigación médica y en la *medicina basada en la evidencia*. Nosotros estamos de acuerdo con Harden [83] cuando afirma que no deberíamos descartar este enfoque con la excusa de la falta de evidencia o de su mayor complejidad. Para muchos autores la respuesta a la pregunta de si esta justificado el enfoque de buscar la mejor evidencia en educación, o lo que es lo mismo si es necesario investigar en el campo de la educación médica, es claramente positiva. Existe la necesidad de que los resultados de la enseñanza sean comprensibles y medibles y una razón primordial es que la investigación en educación médica debe jugar un papel clave en la mejoría de la calidad de la medicina y la asistencia sanitaria. El desarrollo de la *Mejor Evidencia en Educación Médica* constituye una oportunidad de informar, guiar y mantener la investigación en educación médica [227]. En este sentido el diseño de estudios controlados y randomizados como el nuestro es una vía, aunque no la única, de proporcionar buena evidencia en educación médica [85].

De nuestros resultados se deduce que hemos de ser más críticos y menos dogmáticos frente a las ventajas e inconvenientes de determinadas estrategias docentes. La metodología docente ha de estar supeditada a los criterios de la praxis profesional para que la información resultante sea de calidad. En el campo de la Anestesiología es muy importante identificar cuales son las necesidades de aprendizaje de los especialistas y qué factores motivan cambios en su práctica clínica. Sería interesante estudiar la eficacia de la *clase magistral* y el *aprendizaje basado en la discusión de casos/problema* en otras áreas de interés de la Anestesiología, en la valoración de las habilidades técnicas, aptitudes y competencia profesional.

Un objetivo para el futuro sería diseñar evaluaciones que midan un mayor número de resultados o más de un nivel de competencias. Además, es necesario definir las posibles correlaciones entre nuestras evaluaciones y los tests de respuesta múltiple y/o las evaluaciones de habilidades prácticas. Sería igualmente interesante comparar nuevos métodos de evaluación de los conocimientos, extendiendo el análisis a otras habilidades cognitivas e integrativas y determinando el efecto de nuevos métodos de enseñanza. Son necesarios estudios a largo plazo que determinen si las mejorías en las habilidades analizadas mejoran o no la competencia de los alumnos en la práctica clínica. Por último, faltan estudios sobre la influencia de la enseñanza en Anestesiología en la morbimortalidad, pronóstico o grado de satisfacción de los pacientes. En opinión de Epstein [122], este es el futuro de las evaluaciones en educación médica. Debemos desarrollar la metodología apropiada para esta investigación.

El futuro de la enseñanza de la Anestesiología se está redirigiendo a enfatizar los aspectos prácticos de la profesión [229]. Si sólo prestamos atención a las competencias que podemos medir, corremos el riesgo de perder de vista nuestro trabajo profesional como un todo. La *excelencia* en Anestesiología es un concepto de reciente introducción vinculado al sentido de profesionalidad y nexo de unión de las diferentes competencias profesionales entre sí, principalmente las relacionadas con el comportamiento [205]. Este nuevo enfoque no es ni mejor ni peor que la eficacia *basada en competencias* o en los *resultados*; son enfoques complementarios. Los componentes de *excelencia* de la práctica de la Anestesiología son tácitos, no explícitos; incluyen, por ejemplo, descubrir patrones de pensamiento y actuación (conocimiento tácito) [86], el manejo de tareas, trabajo en equipo, situaciones de alerta y toma de decisiones (habilidades no técnicas) [168] o la responsabilidad, el humanismo, el comportamiento ético las habilidades de comunicación y el sentido del bienestar (profesionalismo) [207]. Las metodologías de investigación cualitativas, más comúnmente asociadas a la sociología y psicología, parecen ser una mejor herramienta para su exploración [84,86,230,231] aunque pueden presentar problemas de fiabilidad [114]. La resolución de casos/problema también se ha referido como un medio posible de medir el profesionalismo [232]. No sabemos que métodos de enseñanza mejoraran la *excelencia* ni las dificultades metodológicas que presentara el diseño de este tipo de estudios; tampoco están definidos los métodos para su evaluación. Sea como fuere, creemos que este es un campo muy interesante de investigación futura.

Durante los últimos años se ha insistido en la necesidad de fomentar la evidencia aplicada a los estudios de educación médica [81,83,227]. Esta recomendación enfatiza la necesidad de poner en marcha estudios rigurosos basados en diseños experimentales para el análisis de la eficacia de la enseñanza. Nuestro diseño de investigación ha intentado seguir este principio. Smits et al. en 2002 [20] y Koh et al. en 2008 [16] definieron cuales eran los criterios de calidad que tenían que cumplimentar los estudios de investigación docente. Creemos que la metodología aplicada por nosotros ha incluido los siguientes criterios de fortaleza: estudio controlado, prospectivo, inclusión de un grupo problema y un grupo control, randomización, evaluación objetiva, repetición de las evaluaciones antes y después de la enseñanza, estudio piloto, tamaño muestral, grupos homogéneos y elevado porcentaje de contestaciones. El criterio que no cumplimos fue el seguimiento de los resultados en el tiempo. La búsqueda de la evidencia ha de tener en cuenta, además de la *calidad*, otros factores [83]: *utilidad y extensión*, creemos que la repetición de los resultados en los tres estudios es un signo de consistencia y demuestra la eficacia y aplicabilidad de los métodos docentes en tres niveles diferentes de formación en Anestesiología; *fortaleza*, reconocemos que es posible mejorar los criterios de evaluación incluyendo un mayor número de *ítems* y ampliando la evaluación a un mayor número de campos o habilidades; sin embargo, hemos aplicado formulas de fiabilidad y capacidad de repetición de las evaluaciones y criterios estrictos en la interpretación de los resultados, no sólo estadísticos sino también de relevancia docente; *objetivos*; hemos definido con claridad el nivel de resultados y de evaluación de la investigación; reconocemos la necesidad de diseñar estudios enfocados en un nivel más alto de calidad de evidencia; *contexto*, hemos descrito con detalle el entorno de los tres estudios; es necesario demostrar la aplicabilidad de nuestro trabajo en contextos diferentes. Independientemente de estos factores de calidad somos conscientes de la

dificultad real de aislar el efecto directo de los métodos de enseñanza. La aplicabilidad de nuestros resultados solo podrá ser validada si otros docentes incorporan esta metodología a sus programas.

La educación requiere un gran esfuerzo de planificación, tiempo y recursos. El diseño de estudios de este tipo presenta considerables dificultades metodológicas y potenciales sesgos intrínsecos [95,226]. A pesar de ello, creemos que la investigación docente en los programas de formación de Anestesiología ha de intentar basarse en la evidencia. En este sentido, nosotros consideramos que este estudio aporta un diseño metodológico aplicable y reproducible, que permite la evaluación objetiva de la eficacia de los métodos de enseñanza de acuerdo a los principios de la *educación basada en la evidencia* [83]. Sin embargo, las limitaciones mencionadas nos hacen conscientes de que la interpretación de nuestros resultados ha de realizarse con cautela y no podemos generalizar nuestras observaciones a otros temas de los programas de formación en Anestesiología. La aplicabilidad de nuestra metodología tiene que confirmarse en nuevos estudios para poder comparar nuestros resultados y contrastarlos en diferentes poblaciones de alumnos. Presentamos nuestros resultados con la esperanza de que puedan servir de ayuda a otros profesionales a la hora de decidir las estrategias docentes que a su juicio consideran más apropiadas para los alumnos de Anestesiología en los tres niveles de formación; al mismo tiempo, esperamos que este trabajo sea un estímulo para seguir investigando en el campo de la educación médica en Anestesiología.

La investigación en educación médica es, en opinión de algunos autores, la más difícil de realizar [233]. Su complejidad y las limitaciones innatas que conlleva recomiendan que la investigación se realice a múltiples niveles [227]. En investigación docente, la suma de resultados pequeños es probablemente la fórmula para mejorar la calidad de la enseñanza [216]; creemos que en el campo de la Anestesiología esto no es una excepción.

La adquisición de conocimiento, la capacidad de aplicar el conocimiento para la solución de problemas, el manejo de la información, usar el razonamiento clínico o unir el conocimiento básico y el clínico son competencias que necesitan enfatizarse y practicarse durante la enseñanza. Nuestros resultados demuestran que estas habilidades pueden ser enseñadas eficazmente a los alumnos de Anestesiología. La implementación en los programas de formación de Anestesiología de la *clase magistral* y el *aprendizaje basado en la discusión de casos/problema* tienen ambas un impacto positivo sobre estas competencias.

7.CONCLUSIONES

- 1.- La eficacia de la *clase magistral* y el *aprendizaje basado en la discusión de casos/problema* fue similar en términos de mejoría inmediata de la adquisición de conocimientos, la capacidad de aplicar el conocimiento para la solución de problemas, el manejo de la información, el razonamiento clínico o unir el conocimiento básico y clínico después de una única sesión docente sobre *soporte vital básico, valoración pre-anestésica y embolismo aéreo* en anestesia en los tres niveles de formación en Anestesiología: *pregrado, postgrado y Formación Médica Continuada*.
- 2.- Ambos métodos fueron válidos y aplicables en estos contextos docentes.
- 3.- El método de *evaluación basado en casos*, estructurado por campos de conocimientos que incorporan *ítems* y con evaluaciones antes y después de la enseñanza permitió una valoración objetiva de las habilidades cognitivas e integrativas adquiridas y fortaleció la metodología y la calidad del estudio.
- 4.- El método *clase magistral* utilizado en el presente estudio, estructurado en función de unos objetivos docentes previamente definidos, con una presentación visual apoyada en soporte Power Point o grabación de video, incluyendo esquemas, diagramas o algoritmos, presentado por un profesorado con experiencia docente y clínica en el tema y que incorpore 10 ó 15 minutos de discusión final es eficaz en mejorar las habilidades cognitivas e integrativas analizadas aplicando una evaluación basada en casos.
- 5.- El método novedoso *aprendizaje basado en la discusión de casos problema* resultó viable, su adopción pudo realizarse con los mismos recursos educativos, no implicó cambios en los programas docentes, permitió cumplir los objetivos fijados y fue igualmente eficaz en mejorar las habilidades cognitivas e integrativas analizadas aplicando una evaluación basada en casos. Nuestros resultados justifican la posibilidad de su empleo para la enseñanza de la Anestesiología en *pregrado, postgrado y formación continuada*.
- 6.- El *aprendizaje basado en la discusión de casos/problema* no presentó desventajas para los alumnos en la evaluación de habilidades cognitivas e integrativas basada en casos/problema
- 7.- Tanto la *clase magistral* como el *aprendizaje basado en la discusión de casos/problema* son estrategias apropiadas y útiles para la formación en Anestesiología. La decisión de que método adoptar depende de múltiples factores externos al proceso de enseñanza difíciles de controlar. La demostración de viabilidad y eficacia de los métodos docentes puede ayudar a su incorporación en los programas de formación.
- 8.- Consideramos que el diseño metodológico aplicado en el presente estudio favorece la *investigación basada en la evidencia*.
- 9.- Las limitaciones de nuestro estudio y la revisión de la bibliografía sugieren que son precisos más estudios que confirmen y permitan generalizar nuestros resultados. Consideramos necesario profundizar el diseño de métodos de evaluación que analicen múltiples niveles de resultados y competencias profesionales. La investigación docente es esencial para mejorar la eficacia de la enseñanza en Anestesiología.

Conclusiones

10.- El método tradicional de *clase magistral* y el *aprendizaje basado en la discusión de casos/problema* favorecen por igual las habilidades relacionadas con la aplicación práctica del conocimiento. Los dos métodos no son excluyentes entre sí, ambos tienen su propio lugar en la enseñanza de la Anestesiología.

8.BIBLIOGRAFÍA

Bibliografía

1. McKeachie WJ. Teaching tips, 9th ed. New York: Longwood; 1994.
2. Shakarian DC. Beyond lecture: strategies that work. *J Phys Educ Recreation Dance* 1995; 66: 21-4.
3. Brown G, Manogue M. AMEE Medical Education Guide No. 22: Refreshing lecturing: a guide for lecturers. *Med Teach* 2001; 23: 231-44.
4. Ebert-May D, Brewer CA, Allred S. Innovation in large lectures- teaching for active learning through inquiry. *BioScience* 1997; 47: 601-7.
5. Bligh DA. What's the use of lectures? San Francisco: Jossey-Bass; 2000.
6. Willenkin RL. Lectures in anesthesia training. *Anesth Analg* 1992; 74: 1-2.
7. Landers DF, Becker GL, Newland MC, Peters KR. Lecture practices in United States anesthesiology residencies. *Anesth Analg* 1992; 74: 112-5.
8. Panel on the General Professional Education of the Physician and College Preparation for Medicine: Physicians for the twenty-first century. The GPEP report. *J Med Educ* 1984; 59: 1-208.
9. Norman G. What's the active ingredient in active learning? *Adv Health Sci Educ Theory Pract* 2004; 9: 1-3.
10. Tomey AM. Learning with cases. *J Contin Educ Nurs* 2003; 34: 34-8.
11. Pallie W, Carr DH. The McMaster Medical education philosophy in theory, practice and historical perspective. *Med Teach* 1987; 9: 59-71.
12. Neville AJ, Norman GR. PBL in the undergraduate MD program at McMaster university: three iterations in three decades. *Acad Med* 2007; 82: 370-4.
13. Spencer JA, Jordan RK. Learner centred approaches in medical education. *BMJ* 1999; 318: 1280-3.
14. Kinkade S. A snapshot of the status of problem-based learning in U. S. medical schools, 2003-04. *Acad Med* 2005; 80: 300-1.
15. Baños JE. Aprendizaje basado en problemas y trampa decimológica: dime qué evalúas y te diré que encontrarás. *Rev Esp Anestesiología Reanimación* 2008; 55: 135-6.
16. Koh GC, Khoo HE, Wong ML, Koh D. The effects of problem-based learning during medical school on physician competency: a systematic review. *CMAJ* 2008; 178: 34-41.
17. Williams SM, Beattie HJ. Problem based learning in the clinical setting--a systematic review. *Nurse Educ Today* 2008; 28: 146-54.
18. Branda LA. El aprendizaje basado en problemas. De herejía artificial a *res popularis*. *Educ Med* 2009; 12: 11-23.
19. Dolmans DH, De Grave W, Wolfhagen IH, van der Vleuten CP. Problem-based learning: future challenges for educational practice and research. *Med Educ* 2005; 39: 732-41.
20. Smits PB, Verbeek JH, de Buissonjé CD. Problem based learning in continuing medical education: a review of controlled evaluation studies. *BMJ* 2002; 324: 153-6.
21. Srinivasan M, Wilkes M, Stevenson F, Nguyen T, Slavin S. Comparing problem-based learning with case-based learning: effects of a major curricular shift at two institutions. *Acad Med* 2007; 82: 74-82.
22. Davis MH, Harden RM. AMEE medical education guide nº 15: problem based learning: a practical guide. *Med Teach* 1999; 21: 130-40.
23. Neufeld VR, Barrows HS. The "McMaster Philosophy": an approach to medical education. *J Med Educ* 1974; 49: 1040-50.
24. Heale J, Davis D, Norman G, Woodward C, Neufeld V, Dodd P. A randomized controlled trial assessing the impact of problem-based versus didactic teaching methods in CME. *Res Med Educ* 1988; 27: 72-7.
25. Barrows HS. A taxonomy of problem-based learning methods. *Med Educ* 1986; 20: 481-6.
26. Norman GR, Schmidt HG. The psychological basis of problem-based learning: a review of evidence. *Acad Med* 1992; 67: 557-65.
27. Schmidt HG. Foundation of problem-based learning: Some explanatory notes. *Med Educ* 1993; 27: 422-32.

Bibliografía

28. Maudsley G. Do we all mean the same thing by "problem-based learning"? A review of the concepts and a formulation of the ground rules. *Acad Med* 1999; 74: 178-85.
29. Jones RW. Problem-based learning: description, advantages, disadvantages, scenarios and facilitation. *Anaesth Intensive Care* 2006; 34: 485-8.
30. O'Neill PA, Willis SC, Jones A. A model of how students link problem-based learning with clinical experience through "elaboration". *Acad Med* 2002; 77: 552-61.
31. Rothman AI. Problem-based learning--time to move forward? *Med Educ* 2000; 34: 509-10.
32. Shanley PF. Viewpoint: leaving the "empty glass" of problem-based learning behind: new assumptions and a revised model for case study in preclinical medical education. *Acad Med* 2007; 82: 479-85.
33. Lucas M, García Guasch R, Moret E, Llasera R, Melero A, Canet J. El aprendizaje basado en problemas aplicado a la asignatura de pregrado de Anestesiología, Reanimación y Terapéutica del Dolor. *Rev Esp Anesthesiol Reanim* 2006; 53: 419-25.
34. Des Marchais JE. From traditional to problem-based curriculum: how the switch was made at Sherbrooke, Canada. *Lancet* 1991; 338: 234-7.
35. Des Marchais JE, Bureau MA, Dumais B, Pigeon G. From traditional to problem-based learning: a case report of complete curriculum reform. *Med Educ* 1992; 26: 190-9.
36. Donner RS, Bickley H. Problem-based learning: an assessment of its feasibility and cost. *Hum Pathol* 1990; 21: 881-5.
37. Hendry GD, Ryan G, Harris J. Group problems in problem-based learning. *Med Teach* 2003; 25: 609-16.
38. Dornan T, Hadfield J, Brown M, Boshuizen H, Scherpbier A. How can medical students learn in a self-directed way in the clinical environment? Design-based research. *Med Edu* 2005; 39: 356-64.
39. Hitchcock MA, Anderson AS. Dealing with dysfunctional tutorial groups. *Teach Learn Med* 1997; 9: 19-24;
40. De Grave WS, Dolmans DH, van Der Vleuten CP. Student perceptions about the occurrence of critical incidents in tutorial groups. *Med Teach* 2001; 23: 49-54.
41. Cobb P. Theories of knowledge and instructional design: a response to Colliver. *Teach Learn Med* 2002; 14: 52-5.
42. Schmidt HG. Resolving inconsistencies in tutor expertise research: does lack of structure cause students to seek tutor guidance? *Acad Med* 1994; 69: 656-62.
43. Roberts C, Lawson M, Newble D, Self A, Chan P. The introduction of large class problem-based learning into an undergraduate medical curriculum: an evaluation. *Med Teach* 2005; 27: 527-33.
44. De Grave WS, Dolmans DH, van der Vleuten CP. Profiles of effective tutors in problem-based learning: scaffolding student learning. *Med Educ* 1999; 33: 901-6.
45. Hay PJ, Katsikitis M. The 'expert' in problem-based and case-based learning: necessary or not? *Med Educ* 2001; 35: 22-6.
46. Albanese MA, Mitchell S. Problem-based learning: a review of literature on its outcomes and implementation issues. *Acad Med* 1993; 68: 52-81.
47. Vernon DTA, Blake RL. Does problem-based learning work? A metaanalysis of evaluative research. *Acad Med* 1993; 68: 550-63.
48. Distlehorst LH, Dawson E, Robbs RS, Barrows HS. Problem-based learning outcomes: the glass half-full. *Acad Med* 2005; 80: 294-9.
49. Colliver JA. Educational theory and medical education practice: a cautionary note for medical school faculty. *Acad Med* 2002; 77: 1217-20.
50. Smits PB, de Buissonje CD, Verbeek JH, van Dijk FJ, Metz JC, ten Cate OJ. Problem-based learning versus lecture-based learning in postgraduate medical education. *Scand J Work Environ Health* 2003; 29: 280-7.
51. Barrows H, Tambly R, eds. Problem based learning. New York: Springer-Verlag; 1980.

Bibliografía

52. Savin-Baden M. Critical perspectives on problem-based learning. En: *Problem-Based Learning in Higher Education: Untold Stories*. Buckingham, UK: Open University Press; 2000:124–34.
53. Liu PL. 1996 Problem-based learning discussions program progress to new pinnacles. *ASA newsletter* [serie en Internet] 1996 Julio; 60 (7): [3 páginas].
Disponble en: URL; http://www.asahq.org/Newsletters/1996/07_96/Article10.htm
54. Rosenblatt MA. The educational effectiveness of problem-based learning discussions as evaluated by learner-assessed satisfaction and practice change. *J Clin Anesth* 2004; 16: 596-601.
55. Wenk M, Waurick R, Schotes D, Wenk M, Gerdes C, Van Aken HK, et al. Simulation-based medical education is no better than problem-based discussions and induces misjudgment in self-assessment. *Adv Health Sci Educ Theory Pract* 2009; 14: 159-71.
56. Williams B. Case based learning--a review of the literature: is there scope for this educational paradigm in prehospital education? *Emerg Med J* 2005; 22: 577-81.
57. Richards PS, Inglehart MR. An interdisciplinary approach to case-based teaching: does it create patient-centered and culturally sensitive providers? *J Dent Educ* 2006; 70: 284-91.
58. Tärnvik A. Advantages of using the multiple case method at the clinical stage of medical education. *Med Teach* 2002; 24: 396-401.
59. Smith S, Fryer-Edwards K, Diekema DS, Braddock CH III. Finding effective strategies for teaching ethics: a comparison trial of two interventions. *Acad Med* 2004; 79: 265-71.
60. Marantz PR, Burton W, Steiner-Grossman P. Using the case-discussion method to teach epidemiology and biostatistics. *Acad Med* 2003; 78: 365-71.
61. Johnson JP, Mighten A. A comparison of teaching strategies: lecture notes combined with structured group discussion versus lecture only. *J Nurs Educ* 2005; 44: 319-22.
62. Tärnvik A. Revival of the case method: a way to retain student-centred learning in a post-PBL era. *Med Teach* 2007; 29: e32-e36.
63. Sperle DH. The case method technique in professional training. New York: Teachers College, Columbia University; 1933.
64. Engelberg J. Complex medical case histories as portals to medical practice and integrative, scientific thought. *Am J Physiol* 1992; 263: S45–S54.
65. Thomas RE. Methods of teaching medicine using cases. *Med Teach* 1993; 15: 27–34.
66. Harris NL. Case records of the Massachusetts General Hospital--continuing to learn from the patient. *N Engl J Med* 2003; 348: 2252-4.
67. Schmidt H. Assumptions underlying self-directed learning may be false. *Med Educ* 2000; 34: 243-5.
68. Garvey T, O'Sullivan M, Blake M. Multidisciplinary case-based learning for undergraduate students. *Eur J Dent Educ* 2000; 4: 165-8.
69. Dolmans D, Snellen-Balendong H, Wolfhagen I, van der Vleuten C. Seven principles of effective case design for a problema-based curriculum. *Med Teach* 1997; 19: 185-9.
70. Bonwell CC, Eison JA. *Active learning: creating excitement in the classroom*. Washington DC: Clearinghouse on Higher Education, George Washington University; 1991.
71. Fanning RM, Gaba DM. The role of debriefing in simulation-based learning. *Simul Healthc* 2007; 2: 115-25.
72. Cannon-Bowers JA. Recent advances in scenario-based training for medical education. *Curr Opin Anaesthesiol* 2008; 21: 784-9.
73. Veloski J, Boex JR, Grasberger MJ, Evans A, Wolfson DB. Systematic review of the literatura on assessment, feedback and physicians' clinical performance. BEME Guide nº 7. *Med Teach* 2006; 28: 117-28.
74. Barnes LB, Christensen CR, Hansen AB. *Teaching and the Case Method*. Boston, Massachusetts: Harvard Business School Press; 1994.

Bibliografía

75. Erskine J, Leenders MR, Maufette-Leenders LA. Teaching with Cases. En: Richard Ivey School of Business. The University of Western Ontario. Canada; 1998.
76. Wahlgren CF, Edelbring S, Fors U, Hindbeck H, StÅhle M. Evaluation of an interactive case simulation system in dermatology and venereology for medical students. *BMC Med Educ* 2006; 6: 40.
77. Sandstrom S. Use of case studies to teach diabetes and other chronic illness to nursing students. *J Nurs Educ* 2006; 45: 229-32.
78. Albanese M. Problem-based learning: why curricula are likely to show little effect on knowledge and clinical skills. *Med Educ* 2000; 34: 729-38.
79. Dowd SB, Davidhizar R. Using case studies to teach clinical problem-solving. *Nurse Educ* 1999; 24: 42-6.
80. Allen KL, More FG. Clinical simulation and foundation skills: an integrated multidisciplinary approach to teaching. *J Dent Educ* 2004; 68: 468-74.
81. Van Der Vleuten CP, Dolmans D, Scherpbier A. The need for evidence in education. *Med Teach* 2000; 22: 246-50.
82. Best Evidence Medical Education (BEME): report of meeting - 3-5 December 1999, London, UK. *Med Teach* 2000; 22: 242-5.
83. Harden RM, Grant J, Buckley G, Hart IR. Best Evidence Medical Education. *Adv Health Sci Educ Theory Pract* 2000; 5: 71-90.
84. Malterud K: Qualitative research: Standards, challenges, and guidelines. *Lancet* 2001; 358: 483-8.
85. Norman G. The end fo educational science? *Adv Health Sci Educ* 2008; 13: 385-9.
86. Larsson J: Studying tacit knowledge in anesthesiology: A role for qualitative research. *Anesthesiology* 2009; 110: 443-4.
87. Hart IR. Best evidence medical education. *Med Teach* 1999; 21: 453-4.
88. Wass V, van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. *Lancet* 2001; 357: 945-9.
89. Cate OT. What happens to the student? The neglected variable in education outcomes research. *Adv Health Sci Educ* 2001; 6: 81-8.
90. Harden RM. Outcome-based education: the future is today. *Med Teach* 2007; 29: 625-9.
91. Harden RM. Outcome-based education--the ostrich, the peacock and the beaver. *Med Teach* 2007; 29: 666-71.
92. Harden RM. Developments in outcome-based education. *Med Teach* 2002; 24: 117-20.
93. Kirkpatrick DL. Evaluating training programs: The four levels. San Francisco; Berrett-Koehler Publishers; 1994.
94. Freeth D, Hammick M, Koppel I, Reeves S, Barr H. A critical review of evaluations of interprofessional education. Higher Education Academy Learning and Teaching Support Network for Health Sciences and Practice. Londres; 2003.
Disponibile en: URL; <http://www.health.ltsn.ac.uk/publications/occasionalpaper/occp5.pdf>.
95. Steinert Y, Mann K, Centeno A, Dolmans D, Spencer J, Gelula M, et al. A systematic review of faculty development initiatives designed to improve teaching effectiveness in medical education: BEME Guide No. 8. *Med Teach* 2006; 28: 497-526.
96. Albino JE, Young SK, Neumann LM, Kramer GA, Andrieu SC, Henson L, et al. Assessing dental students' competence: best practice recommendations in the performance assessment literature and investigation of current practices in predoctoral dental education. *J Dent Educ* 2008; 72: 1405-35.
97. Kassirer JP. Pseudoaccountability. *Ann Intern Med* 2001; 134: 587-90.
98. Talbot M: Monkey see, monkey do: A critique of the competency model in graduate medical education. *Med Educ* 2004; 38: 580-1.
99. Tetzlaff JE. Assessment of competency in anesthesiology. *Anesthesiology* 2007; 106: 812-25.

Bibliografía

100. Norman GR. Defining competence: a methodological review. En: Neufeld VR, Norman GR (Eds.). *Assessing clinical competence*. New York: Springer; 1985: 15-35.
101. Martínez-Carretero JM. Los métodos de evaluación de la competencia profesional: la evaluación clínica objetiva estructurada (ECOPE). *Educ med* [serie en Internet], 2005 Septiembre; 8 (Supl 2):18-22.
Disponible en: URL; http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1575-18132005000600007&lng=es&nrm=iso
102. Proyecto Tuning. Informe final. Bilbao: Universidad de Deusto; 2003.
103. Kane MT. The assessment of professional competence. *Eval Health Prof* 1992; 15: 163-82.
104. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA* 2002; 287: 226-35.
105. Gomar C, Montero A, Puig M, Sabaté A, Fábregas N, Valero R, et al. Competencias en Anestesiología, Reanimación y Terapéutica del Dolor de un graduado de Medicina. Consenso del profesorado de las Universidades de Cataluña. *Rev Esp Anestesiol Reanim* 2009 (en prensa).
106. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990; 65: S63-7.
107. Greaves JD, Grant J. Watching anaesthetists work: using the professional judgement of consultants to assess the developing clinical competence of trainees. *Br J Anaesth* 2000; 84: 525-33.
108. Leach DC. Competence is a habit. *JAMA* 2002; 287: 243-4.
109. Bloom BS, Englehart MD, Furst EJ, Hill WH, Krathwohl DR. *Taxonomy of educational objectives. Handbook 1: cognitive domain*. New York: Longmans; 1956.
110. Kearney RA, Puchalski SA, Yang HY, Skakun EN. The inter-rater and intra-rater reliability of a new Canadian oral examination format in anesthesia is fair to good. *Can J Anaesth* 2002; 49: 232-6.
111. Van der Vleuten C. Validity of final examinations in undergraduate medical training. *BMJ* 2000; 321: 1217-9.
112. Schuwirth LW, Southgate L, Page GG, Paget NS, Lescop JM, Lew SR, et al. When enough is enough: a conceptual basis for fair and defensible practice performance assessment. *Med Educ* 2002; 36: 925-30.
113. Harden RM, Gleeson FA. ASME medical educational booklet nº 8: assessment of medical competence using an objective structured clinical examination (OSCE). *J Med Educ* 1979; 13: 41-54.
114. Newble DI. Assessment of clinical competence. *Br J Anaesth* 2000; 84: 432-3.
115. Boulet JR, Murray D, Kras J, Woodhouse J, McAllister J, Ziv A. Reliability and validity of a simulation-based acute care skills assessment for medical students and residents. *Anesthesiology* 2003; 99: 1270-80.
116. Yee B, Naik VN, Joo HS, Savoldelli GL, Chung DY, Houston PL, et al. Nontechnical skills in anesthesia crisis management with repeated exposure to simulation-based education. *Anesthesiology* 2005; 103: 241-8.
117. Van der Vleuten CP, Norman GR, De Graaff E. Pitfalls in the pursuit of objectivity: issues of reliability. *Med Educ* 1991; 25: 110-8.
118. Mases A, Bermejo S, Mojal S, Rojo A, Escolano F. Development and validation of a questionnaire to evaluate competent practice in anaesthesiology residents. *Education, research and presentation* 2008: 226 (abstract).
119. Rhoton MF, Barnes A, Flashburg M, Ronai A, Springman S. Influence of anesthesiology residents' noncognitive skills on the occurrence of critical incidents and the residents' overall clinical performances. *Acad Med* 1991; 66: 359-61.
120. Lockyer J. Multisource feedback in the assessment of physician competencies. *J Contin Educ Health Prof* 2003; 23: 4-12.
121. Jahangiri L, Mucciolo TW, Choi M, Spielman AI. Assessment of teaching effectiveness in U.S. Dental schools and the value of triangulation. *J Dent Educ* 2008; 72: 707-18.
122. Epstein RM. Assessment in medical education. *N Engl J Med* 2007; 25: 387-96.
123. Friedman Ben David M, Davis MH, Harden RM, Howie PW, Ker J, Pippard MJ. AMEE Medical Education Guide No. 24: Portfolios as a method of student assessment. *Med Teach* 2001; 23:535-51.

Bibliografía

124. Shumway JM, Harden RM; Association for Medical Education in Europe. AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. *Med Teach* 2003; 25: 569-84.
125. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teach* 2007; 29: 855-71.
126. Bandaranayake RC. Setting and maintaining standards in multiple choice examinations: AMEE Guide No. 37. *Med Teach* 2008; 30: 836-45.
127. Schuwirth LW, Verheggen MM, van der Vleuten CP, Boshuizen HP, Dinant GJ. Do short cases elicit different thinking processes than factual knowledge questions do? *Med Educ* 2001; 35: 348-56.
128. Cornet A. Nuevos planteamientos didácticos: ¿al innovar en docencia, mejoramos el aprendizaje? *Educación Médica (Barcelona)* [serie en Internet] 2006 Septiembre; 8 (Supl 1): S18. Disponible en: URL; http://scielo.isciii.es/scielo.php?script=sci_issue&pid=1575-1813&lng_es&nrm=iso
129. Reamer LE. Is there an evidence-based approach to anesthesia education? *Best Pract Res Clin Anaesthesiol* 2005; 19: 137-52.
130. Mikail CN, Hearney E, Nemesure B. Increasing physician awareness of the common uses and contraindications of herbal medicines: utility of a case-based tutorial for residents. *J Altern Complement Med* 2003; 9: 571-6.
131. Koles P, Nelson S, Stolfi A, Parmelee D, Destephen D. Active learning in a Year 2 pathology curriculum. *Med Educ* 2005; 39:1045-55.
132. Hammick M. A BEME review: a little illumination. *Med Teach* 2005; 27: 1-3
133. Handley AJ, Koster R, Monsieurs K, Perkins GD, Davies S, Bossaert L. European Resuscitation Council Guidelines for Resuscitation 2005. Section 2. Adult basic life support and use of automated external defibrillators. *Resuscitation* 2005; 67S1: S7 - S23.
134. Baskett PJF, Nolan JP, Handley A, Soar J, Biarent D, Richmond S. European Resuscitation Council Guidelines for Resuscitation 2005: Section 9. Principles of training in resuscitation. *Resuscitation* 2005; 67S1: S181-S189.
135. Jordan T, Bradley P. A survey of basic life support training in various undergraduate health care professions. *Resuscitation* 2000; 47: 321-3.
136. Philips PS, Nolan JP. Training in basic and advanced life support in UK medical schools: questionnaire survey. *Br Med J* 2001; 323: 22-3.
137. Hoke RS, Handley AJ. A reference basic life support provider course for Europe. *Resuscitation* 2006; 69: 413-9.
138. Chamberlain D, Hazinski MF. European Resuscitation Council, American Heart Association, Heart and Stroke Foundation of Canada, Resuscitation Council of Southern Africa, Australia and New Zealand Resuscitation Council, Consejo Latino-Americano de Resucitación. Education in Resuscitation: an ILCOR symposium: Utstein Abbey: Stavanger, Norway: June 22-24, 2001. *Circulation* 2003; 108: 2575-94.
139. Williamson JA, Helps SC, Westhorpe RN, Mackay P. Crisis management during anaesthesia: embolism. *Qual Saf Health Care* 2005; 14: e17.
140. Domaingue CM. Anaesthesia for neurosurgery in the sitting position: a practical approach. *Anaesth Intensive Care* 2005; 33: 323-31.
141. Smith DS, Osborn I. Posterior fossa: anesthetic considerations. En: Cottrell JE, Smith DS, (Eds.). *Anesthesia and Neurosurgery*, 4th ed. St Louis: Mosby; 2001: 335-51.
142. Scherpereel P. Foundation for european education in anaesthesiology (FEEA). *Eur J Anaesthesiol* 2000; 17: 75-6.
143. Hernández C, Villalonga A. Embolia gaseosa en una paciente intervenida de un tumor de fosa posterior. En: Gomar C, Villalonga A (Eds.). *Casos Clínicos Anestesiología II*, 2ª ed. Barcelona: Masson; 2005: 480-8.
144. Marrugat J, Vila J, Pavesi M, Sanz F. Estimation of the sample size in clinical and epidemiological investigations. *Med Clin (Barc)* 1998; 111: 267-76.
145. Altman DG. *Practical statistics for medical research*. New York: Chapman & Hall/CRC; 1991.

Bibliografía

146. Mallampati SR. Clinical signs to predict difficult tracheal intubation (hypothesis). *Can Anaesth Soc J* 1983; 30: 316-17.
147. Wolters U, Wolf T, Stützer H, Schröder T. ASA classification and perioperative variables as predictors of postoperative outcome. *Br J Anaesth* 1996; 77: 217-22.
148. Suarez S, Ornaque I, Fábregas N, Valero R, Carrero E. Venous air embolism during Parkinson surgery in patients with spontaneous ventilation. *Anesth Analg* 1999; 88: 793-4.
149. Tiwari A, Lai P, So M, Yuen K. A comparison of the effects of problem-based learning and lecturing on the development of students critical thinking. *Med Educ* 2006; 40: 547-54.
150. Hwang SY, Kim MJ. A comparison of problem-based learning and lecture-based learning in an adult health nursing course. *Nurse Educ Today* 2006; 26: 315-21.
151. Mellon AF, Mellon J. Logical debate on problem based learning *BMJ* 2006; 332: 550-1.
152. Cohen-Schotanus J, Muijtjens AMM, Shönrock-Adema J, Geertsma J, van der Vlieten CPM. Effects of conventional and problem-based learning on clinical and general competences and career development. *Med Edu* 2008; 42: 256-65.
153. Colliver JA, Markwell SJ. Research on problem-based learning: the need for critical analysis of methods and findings. *Med Educ* 2007; 41: 533-5.
154. Wood DF. Problem based learning. *BMJ* 2008; 336: 971.
155. Engel F, Hendricson W. A case-based learning model in orthodontics. *J Dent Educ* 1994; 58: 762-77.
156. Kassebaum D, Averbach R, Fryer G. Student preference for a case-based vs.lecture instructional format. *J Dent Educ* 1991; 55: 781-4.
157. Pearson T, Barker W, Fisher S, et al. Integration of the case-based series in population-orientated prevention into a problem-based medical curriculum. *Am J Prev Med* 2003; 24: 102-7.
158. Hansen W, Ferguson K, Sipe C, et al. Attitudes of faculty and students toward case-based learning in the third-year obstetrics and gynecology clerkship. *Am J Obstet Gynecol* 2005; 192: 644-7.
159. Ward R. Active, collaborative and case-based learning with computer-based case scenarios. *Comput Educ* 1997; 30: 103-10.
160. Nadiu S, Oliver M, Koronios A. Approaching clinical decision making in nursing practice with interactive multimedia and case-based reasoning. *The Interactive Multimedia Electronic Journal of Computer-enhanced Learning* [serie en Internet] 1999 Octubre;1(2): [11 páginas]. Disponible en: URL: <http://imej.wfu.edu/>
161. Thomas M, O'Connor F, Albert M, Boutain D, Brandt P. Case-based teaching and learning experiences. *Iss Mental Health Nurs* 2001; 22: 517-31.
162. Parker KJ. Lecturing and loving it: applying the information processing method. *The Clearing House* 1993; 67: 8-11.
163. Cleave-Hogg D, Benedict C. Characteristics of good anaesthesia teachers. *Can J Anaesth* 1997; 44: 587-91.
164. Smith A, Goodwin D, Mort M, Pope C. Expertise in practice: an ethnographic study exploring acquisition and use of knowledge in anaesthesia. *Br J Anaesth* 2003; 91: 319-28.
165. Orlander JD, Barber TW, Fincke BG. The morbidity and mortality conference: the delicate nature of learning from error. *Acad Med* 2002; 77: 1001-6.
166. O'Brien MA, Freemantle N, Oxman AD, Wolf F, Davis DA, Herrin J. Continuing education meetings and workshops: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2001, (1): CD003030.
167. Satterlee WG, Eggers RG, Grimes DA. Effective medical education: insights from the Cochrane Library. *Obstet Gynecol Surv* 2008; 63: 329-33.
168. Glavin RJ: Excellence in anesthesiology: The role of nontechnical skills. *Anesthesiology* 2009; 110: 201-3.
169. Parkes J, Hyde C, Deeks J, Milne R. Teaching critical appraisal skills in health care settings. *Cochrane Database Syst Rev* 2001; (3):CD001270.

Bibliografia

170. Karaalp A, Akici A, Kocabasoglu YE, Oktay S. What do graduates think about a two-week rational pharmacotherapy course in the fifth year of medical education? *Med Teach* 2003; 25: 515-21.
171. Dugdale A. Traditional medical education and the new path--they are not mutually exclusive. *Med Educ* 2001; 35: 304.
172. Haidet P, Morgan RO, O'Malley K, Moran BJ, Richards BF. A controlled trial of active versus passive learning strategies in a large group setting. *Adv Health Sci Educ Theory Pract* 2004; 9: 15-27.
173. Stehr SN, Muller M, Frank MD et al. Teaching methods in anesthesia and intensive care medicine. The new legislation and its possibilities for the specialty. *Anaesthetist* 2005; 54: 385-93.
174. Rodrigues de Oliveira G, Schonhorst L. Problem-based learning implementation in an intensive course of anaesthesiology: a preliminary report on residents' cognitive performance and perceptions of the educational environment. *Med Teach* 2005; 27: 382-4.
175. Chang CH, Yang CY, See LC, Lui PW. High satisfaction with problem-based learning for anaesthesia. *Chang Gung Med J* 2004; 27: 654-62.
176. Belfield C, Thomas H, Bullock A, Eynon R, Wall D. Measuring effectiveness for best evidence medical education: a discussion. *Med Teach* 2001; 23: 164-70.
177. Eva KW, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. *Acad Med* 2005; 80 (10 Suppl): S46-54.
178. Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA* 2006; 296: 1094-102.
179. Eva KW, Cunnington JP, Reiter HI, Keane DR, Norman GR. How can I know what I don't know? Poor self assessment in a well-defined domain. *Adv Health Sci Educ Theory Pract* 2004; 9: 211-24.
180. Langendyk V. Not knowing that they do not know: self-assessment accuracy of third-year medical students. *Med Educ* 2006; 40: 173-9.
181. Barrows HS. Innovations without appropriate assessment are of limited usefulness. *Teach Learn Med* 2008; 20: 287.
182. Sibley JC, Sackett DL, Neufeld V, Gerrard B, Rudnick KV, Fraser W. A randomized trial of continuing medical education. *N Eng J Med* 1982; 306: 511-5.
183. Lampe M. Appraisal and reassessment of the specialist in anaesthesia. *Best Pract Res Clin Anaesthesiol* 2002; 16: 391-400.
184. Kramer GA, Albino JE, Andrieu SC, Hendricson WD, Henson L, Horn BD, et al. Dental student assessment toolbox. *J Dent Educ* 2009; 73:12-35.
185. Shin JH, Haynes RB, Johnston ME. Effect of problem-based, self-directed undergraduate education on life-long learning. *CMAJ* 1993; 48: 969-76.
186. Davis P, Kvern B, Donen N, Andrews E, Nixon O. Evaluation of a problem-based learning workshop using pre- and post-test objective structured clinical examinations and standardized patients. *J Contin Educ Health Prof* 2000; 20: 164-70.
187. Ironside PM. Using narrative pedagogy: learning and practising interpretive thinking. *J Adv Nurs* 2006; 55: 478-86.
188. Madden C. Undergraduate nursing students' acquisition and retention of CPR knowledge and skills. *Nurse Educ Today* 2006; 26: 218-27.
189. Reder S, Cummings P, Quan L. Comparison of three instructional methods for teaching cardiopulmonary resuscitation and use of an automatic external defibrillator to high school students. *Resuscitation* 2006; 69: 443-53.
190. Chamberlain D, Smith A, Woollard M, Colquhoun M, Handley AJ, Leaves S, Kern KB. Trials of teaching methods in basic life support (3): comparison of simulated CPR performance after first training and at 6 months, with a note on the value of re-training. *Resuscitation* 2002; 53: 179-87.

Bibliografia

191. Steadman RH, Coates WC, Huang YM, Matevosian R, Larmon BR, McCullough L, Ariel D. Simulation-based training is superior to problem-based learning for the acquisition of critical assessment and management skills. *Crit Care Med* 2006; 34: 151-7.
192. Blum RH, Raemer DB, Carroll JS, Sunder N, Feinstein DM, Cooper JB. Crisis-resource management training for an anaesthesia faculty: a new approach to continuing education. *Med Educ* 2004; 38: 45-55.
193. Weller J, Wilson L, Robinson B. Survey of change in practice following simulation-based training in crisis management. *Anaesthesia* 2003; 58: 471-9.
194. Issenberg SB, McGaghie WC, Hart IR, Mayer JW, Felner JM, Petrusa ER, Waugh RA, Brown DD, Safford RR, Gessner IH, Gordon DL, Ewy GA. Simulation technology for health care professional skills training and assessment. *JAMA* 1999; 282: 861-6.
195. Kurrek MM, Devitt JH. The cost for construction and operation of a simulation centre. *Can J Anaesth* 1997; 44: 1191-5.
196. Cumin D, Merry AF. Simulators for use in anaesthesia. *Anaesthesia* 2007; 62: 151-62.
197. Marinopoulos SS, Dorman T, Ratanawongsa N, Wilson LM, Ashar BH, Magaziner JL, Miller RG, Thomas PA, Prokopowicz GP, Qayyum R, Bass EB. Effectiveness of continuing medical education. *Evid Rep Technol Assess* 2007; 149: 1-69.
198. Wong AK. Full scale computer simulators in anesthesia training and evaluation. *Can J Anaesth* 2004; 51: 455-64.
199. Issenberg SB, McGaghie WC, Petrusa ER, Gordon DL, Scalese RJ. BEME Guide No 4: Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach* 2005; 27:10-28.
200. McFetrich J. A structured literature review on the use of high fidelity patient simulators for teaching in emergency medicine. *Emerg Med J* 2006; 23: 509-11.
201. Lippert A, Nielsen MS, Ostergaard D. Medical simulation in anesthesiology. The Danish Society of Anesthesiology and Intensive Medicine. *Ugeskr Laeger* 2004; 166: 1102.
202. Gardiner J. Simulators in obstetric anaesthesia are a waste of time. *Int J Obstet Anesth* 2006; 15: 44-6.
203. Lohman MC, Finkelstein M.. Designing cases in problem-based learning to foster problem-solving skill. *Eur J Dent Educ* 2002; 6: 121-7.
204. Norman G. Beyond PBL. *Adv Health Sci Educ Theory Pract* 2004; 9: 257-60.
205. Smith A. In search of excellence in anesthesiology. *Anesthesiology* 2009; 110: 4-5.
206. Reich DL, Uysal S, Bodian CA, Gabriele S, Hibbard M, Gordon W, Sliwinski M, Kayne RD. The relationship of cognitive, personality, and academic measures to anesthesiology resident clinical performance. *Anesth Analg* 1999; 88: 1092-100.
207. Tetzlaff JE. Professionalism in anesthesiology. "What is it?" or "I know it when I see it". *Anesthesiology* 2009; 110: 700-2.
208. Whitcomb ME. More on competency-based education. *Acad Med* 2004; 79: 493-4.
209. Herzig S, Linke RM, Marxen B, Borner U, Antepohl W. Long-term follow up of factual knowledge after a single, randomised problem-based learning course. *BMC Med Educ* 2003; 3:3.
210. Moret E, García-Guasch R, Oller B, Rognoni G, Abarca L. Retention and assimilation after one year of problema-based learning of anesthesia. *Education, research and presentation* 2008: 224 (abstract).
211. Beers GW, Bowden S. The effect of teaching method on long-term knowledge retention. *J Nurs Educ* 2005; 44: 511-14.
212. Sundblad G, Sigrell B, Knotsson L, Lindkvist J, Lindkvist C. Students' evaluation of a learning method: a comparison between problem-based learning and more traditional methods in a specialist university training program in psychotherapy. *Med Teach* 2002; 24: 268-72.
213. Leon AC. Implications of clinical trial design on sample size requirements. *Schizophr Bull* 2008; 34: 664-9.

Bibliografía

214. Pals SL, Murray DM, Alfano CM, Shadish WR, Hannan PJ, Baker WL. Individually randomized group treatment trials: a critical appraisal of frequently used design and analytic approaches. *Am J Public Health* 2008; 98: 1418-24.
215. White M, Michaud G, Pachev G, Lirenman D, Kolenc A, FitzGerald JM. Randomized trial of problem-based versus didactic seminars for disseminating evidence-based guidelines on asthma management to primary care physicians. *J Contin Educ Health Prof* 2004; 24: 237-43.
216. Norman GR, Schmidt HG. Effectiveness of problem-based learning curricula: theory, practice and paper darts. *Med Educ* 2000; 34: 721-8.
217. Nolan J. European Resuscitation Council guidelines for resuscitation 2005. Section 1. Introduction. *Resuscitation* 2005; 67S1: S3-S6.
218. Welch WW, Walberg HJ. Pretest and sensitisation effects in curriculum evaluation. *Amer Educ Res J* 1970; 7: 605-14.
219. Smits PB, Verbeek JH, Ten Cate TJ, Metz JC, van Dijk FJ. Factors predictive of successful learning in postgraduate medical education. *Med Edu* 2004; 38: 758-66.
220. Peplow P. Attitudes and examination performance of female and male medical students in an active, case-based learning programme in anatomy *Med Teach* 1998; 20: 349-55
221. Prideaux D. Researching the outcomes of educational interventions: a matter of design. *BMJ* 2002; 324: 126-7.
222. Draper SW. The Hawthorne effect and other expectancy effects: a note [serie en Internet]. 2005 Marzo [15 páginas]. Disponible en: URL: <http://www.psy.gla.ac.uk/~steve/hawth.html>
223. Davis DA, Thomson MA, Oxman AD, Haynes RB. Evidence for the effectiveness of CME. A review of 50 randomized controlled trials. *JAMA* 1992; 268: 1111-7.
224. Davis D, O'Brien MA, Freemantle N, Wolf FM, Mazmanian P, Taylor-Vaisey A. Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? *JAMA* 1999; 282: 867-74.
225. Norman G. Best evidence medical education and the perversity of humans as subjects. *Adv Health Sci Educ Theory Pract* 2001; 6: 1-3.
226. Lechner SK. Evaluation of teaching and learning strategies. *Med Educ Online* [serie en Internet] 2001; 6: 4. Disponible en: URL; <http://www.med-ed-online.org>
227. Dauphinee WD, Wood-Dauphinee S. The need for evidence in medical education: the development of best evidence medical education as an opportunity to inform, guide, and sustain medical education research. *Acad Med* 2004; 79: 925-30.
228. Russell TL. The no significant difference phenomenon. 2008. Disponible en: URL; <http://www.nosignificantdifference.org/>
229. Smith A: Reaching the parts that are hard to reach: Expanding the scope of professional education in anaesthesia. *Br J Anaesth* 2007; 99: 453-6.
230. Fletcher G, Flin R, McGeorge P, Glavin R, Maran N, Patey R: Anaesthetists Non-Technical Skills (ANTS): Evaluation of a behavioural marker system. *Br J Anaesth* 2003; 90: 580-8.
231. Dorotta I, Staszak J, Takla A, Tetzlaff JE: Teaching and evaluating professionalism for anesthesiology residents. *J Clin Anesth* 2006; 18:148-60.
232. Ginsburg S, Regehr G, Hatala R, McNaughton N, Frohna A, Hodges B, et al. Context, conflict, and resolution: a new conceptual framework for evaluating professionalism. *Acad Med* 2000; 75 (10 Suppl): S6-S11.
233. Berliner DC. Educational research: the hardest science of all. *Educ Res* 2002; 36: 110-2.

9.ANEXOS

9.1. Artículo 1.

Carrero E, Gomar C, Penzo W, Fabregas N, Valero R, Sanchez-Etayo G. Teaching basic life support algorithms by either multimedia presentations or case based discussion equally improves the level of cognitive skills of undergraduate medical students. *Med Teach* 2009; Feb 24:e1-e7. Disponible en: URL; <http://dx.doi.org/10.1080/01421590802512896>

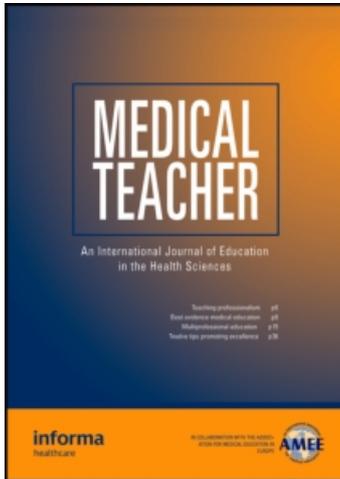
This article was downloaded by: [Consorti de Biblioteques Universitaries de Catalunya]

On: 26 February 2009

Access details: Access Details: [subscription number 789296668]

Publisher Informa Healthcare

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Medical Teacher

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title-content=t713438241>

Teaching basic life support algorithms by either multimedia presentations or case based discussion equally improves the level of cognitive skills of undergraduate medical students

Enrique Carrero ^a; Carmen Gomar ^a; Wilma Penzo ^a; Neus Fábregas ^a; Ricard Valero ^a; Gerard Sánchez-Etayo ^a

^a University of Barcelona, Spain

First Published on: 24 February 2009

To cite this Article Carrero, Enrique, Gomar, Carmen, Penzo, Wilma, Fábregas, Neus, Valero, Ricard and Sánchez-Etayo, Gerard(2009)'Teaching basic life support algorithms by either multimedia presentations or case based discussion equally improves the level of cognitive skills of undergraduate medical students',Medical Teacher,

To link to this Article: DOI: 10.1080/01421590802512896

URL: <http://dx.doi.org/10.1080/01421590802512896>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

WEB PAPER

Teaching basic life support algorithms by either multimedia presentations or case based discussion equally improves the level of cognitive skills of undergraduate medical students

ENRIQUE CARRERO, CARMEN GOMAR, WILMA PENZO, NEUS FÁBREGAS, RICARD VALERO & GERARD SÁNCHEZ-ETAYO

University of Barcelona, Spain

Abstract

Background: There is no evidence on the best method for teaching Basic Life Support (BLS).

Aims: To compare two methods for teaching BLS, assessing the level of cognitive skills.

Methods: Randomized, prospective study including 68 medical students. BLS algorithms were taught for 60 minutes using either a multimedia presentation (Group I, $n=34$) or case based discussion (Group II, $n=34$). Assessments included a scenario-based quiz test and an error-pinpointing video, which the students completed before (T1) and after (T2) teaching. Comparisons between both groups were made on scores of the assessments, actual increases in scores (final value – initial value) and score gains (actual increase/potential increase).

Results: No significant differences were found between the groups in any of the recorded scores. Both groups improved their T2 scores (p values <0.001). The actual increases in scores and the score gains were similar in both groups. Test scores improved in 55.9% of students in Group I and 58.8% in Group II; video scores improved in 85.3% of Group I and in 82.3% of Group II.

Conclusions: BLS teaching by either multimedia presentations or case based discussion equally improves the level of cognitive skills among medical students.

Introduction

The primary objectives of teaching basic life support (BLS) are learning the sequence of action according to international guidelines and acquiring the necessary skills for its proper application (Handley et al. 2005; Baskett et al. 2005).

Despite efforts to prepare and update BLS guidelines, recommendations on the best instructional method are still needed (Jordan & Bradley 2000; Philips & Nolan 2001); a few models have been proposed recently for training non-healthcare professionals (Hoke & Handley 2006). In order to establish significant levels of evidence, the International Liaison Committee on Resuscitation (ILCOR) has adopted the goal of designing randomized controlled studies in educational research to compare standard teacher-based instructional methods versus newer methods focused on the student's participation process (Chamberlain & Hazinski 2003).

In our Medical School, BLS teaching is included in the curriculum as a 2.5-hour-long professor-led station. Action algorithms are taught for one hour using an instructional home-made videotape plus a Power Point presentation; for the remaining hour and a half, students use manikins to practice the BLS sequence of action, airway management,

Practice points

- Instructional methodology in BLS should be based on evidence.
- The best method to teach Universal BLS Algorithms is not defined.
- BLS teaching by either multimedia or case based discussion equally improved cognitive skills among medical students.
- An objective method to assess BLS instruction efficacy is described.
- Blind studies with sufficient samples in the field are difficult.

ventilation, external chest compression and automated external defibrillation. The students' assessment is theoretical and is included as multiple-choice test questions in the final examination of the course. Following the ILCOR's recommendations, we have included the discussion of case scenarios in this one-hour session and we have focused the assessment on aspects related to the application of knowledge.

The aim of this study was to assess if the teaching method based on discussion of case scenarios simulated on manikins

Correspondence: Enrique Carrero, Department of Anaesthesiology, Hospital Clínic, Villarroel 170, 08036 Barcelona, Spain. Tel: +34 932275558; fax: +34 93 4517553; email: ecarrero@clinic.ub.es

by the teacher improves the students capacity to apply acquired knowledge on BLS compared to our traditional teaching method based on multimedia of BLS.

Methods

Subjects

To determine the sample size, the average efficacy was estimated at 60%; we considered a difference in scoring to be relevant at 25%; we established an α risk = 0.05, a statistical power $(1 - \beta) = 0.8$ and we estimated a case loss of about 5%. The resulting sample size was 35 students in each study group. The study population included 70 students in their third year of Medical School at the University of Barcelona, Spain, who attended four consecutive stations on BLS. Participation in the study was voluntary and confidential and had no influence on the academic scoring. All students included in the final analysis (Figure 1) signed the informed consent form. We recorded the students' previous CPR training experience. We excluded those students who did not complete all the phases of the study. Figure 2 shows a schematic representation of the study protocol.

Teaching

The station dealt with BLS according to the European Resuscitation Council (ERC) guidelines (Handley et al. 2005). The two participating professors were certified ERC BLS instructors with more than 5 years of BLS teaching experience and agreed upon the aims of the station before the study. All the students were taught BLS algorithms in a 60-minute session.

Closed envelopes containing a number between 1 and 70 were randomly distributed among medical students (Figures 1 and 2). Students were randomized by a computer-generated procedure (two blocking restriction). Medical students and instructors were not blinded to the educational method applied.

Students in Group I received the content on a non-interactive multimedia format: 20 minutes home-made videotape on BLS, plus 25 minutes of a Power Point presentation and 15 minutes for questions and answers. The videotape and Power Point presentation contents were complementary. The videotape included simulated BLS outside the hospital performed by one or two rescuers with all the ILCOR algorithms steps: victim detection, both victim and rescuer safety, consciousness assessment, help request, upper airway patency, phone call to 112, chest compressions, sequence chest compressions/ventilations and victim reassessment. The Power Point presentation included the epidemiology, etiology and pathophysiology of the cardiac arrest, available evidences on effectiveness of BLS maneuvers, airway devices, ventilation with bag-valve-mask, safety lateral position and performance in front foreign-body airway obstruction. Four work stations were organized with 8–10 students participating in each one. Teaching resources was a computer for all students in each station and the constant presence of the teacher.

Students in Group II directly started with three case-scenario discussions: (1) a victim with non-traumatic cardio respiratory arrest; (2) an unconscious victim with effective and non-effective spontaneous respiration; and (3) a victim with

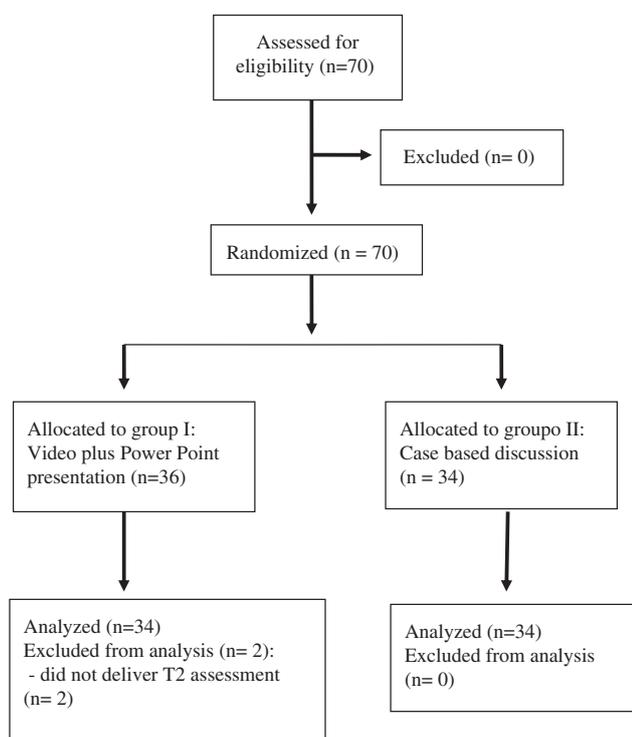


Figure 1. Flow chart showing the medical students' progress in the study.

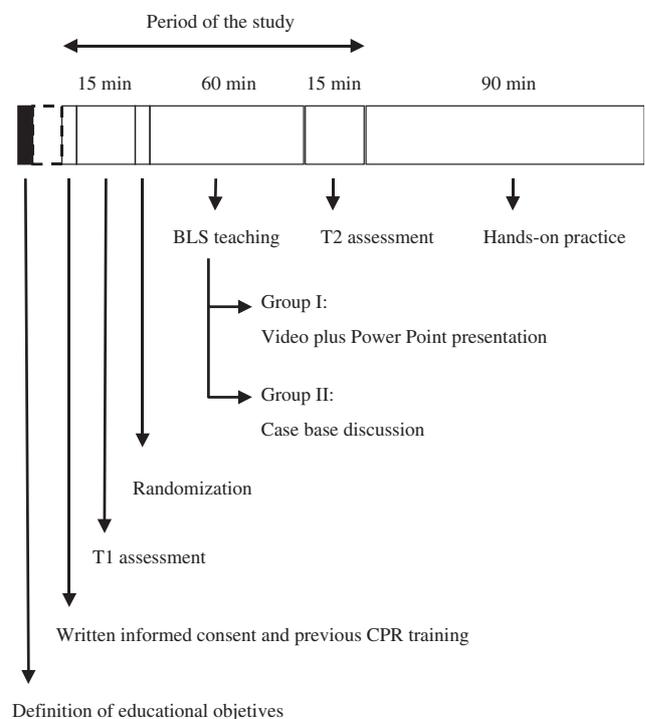


Figure 2. Outline of the methodology followed in the study. T1 assessment: test and video recording before teaching the basic life support (BLS) algorithms; T2 assessment: after teaching the BLS algorithms; CPR: cardiopulmonary resuscitation.

foreign-body airway obstruction. Four work stations were organized with 8–10 students participating in each one. Students did not receive the cases in advance to prepare them. The teacher performed the actions on the manikin during the case scenarios simulation, interacting with the students. The discussion was led by the teacher who detected and corrected the gaps and misunderstandings. The teacher was a subject-matter expert and had experienced the cases in reality. Resources included one Resusci[®] Anne SkillReporter[™] manikin of BLS (Laerdal[®]), barrier devices for mouth-to-mouth ventilation, oropharyngeal cannulae, and one bag–valve–mask system. To simulate safety lateral position and maneuvers for foreign body airway obstruction, the teacher asked for voluntary collaboration of one student. The teaching objectives for both groups in all the stations were the same.

Following the multimedia presentation in Group I and case based discussion in Group II, the station was completed in both groups with a 90-minute hands-on practice on manikins using the BLS equipment. The teachers redressed the students during the hands-on practice. The study objective did not include the effect from the hands-on practice on cognitive skills because this teaching session could have result in a ceiling effect, eliminating the potential differences between the two teaching methods.

Assessments

Assessments included a test and a home-made videotape (Appendix 1). The test included three open-answer questions about three clinical scenarios. The students were asked to answer which BLS maneuver was indicated in each case. The video footage showed four wrong actions in the execution of the BLS sequence on a cardiorespiratory arrest performed by one rescuer on an adult manikin (Resusci[®] Anne). The students were asked to identify the incorrect actions and list them on a template. Assessments did not include hands-on skills.

A pilot trial was performed 6 months earlier for the design of this study. Twenty-nine students at 2 stations on BLS completed the assessments, the test and the video, and were scored by the same two evaluators who would later score the participants in our study. The evaluators confirmed that the assessments were understandable and could be completed within 15 minutes. This pilot trial was used to exclude the test and video data that led to disagreement between the evaluators and to prepare the final design for the test and the video.

Assessments were completed before (T1) and immediately after (T2) the application of the two methods for teaching BLS algorithms (Figure 2). T2 assessment was completed before the 90 min hands-on practice to avoid the possible influence of the latter in the results. Students were given no more than 15 minutes to complete the assessments and were not allowed to make comments on the assessments until finishing T2 assessment.

The primary endpoints assessed were two different fields of cognitive skills in BLS: field 1 ‘decisions based on BLS sequence’ (test) and field 2 ‘detection of errors on BLS performance’ (video recording). We counted the total number

of correct answers at T1 and T2 assessments. Each correct answer scored 1 point, whereas wrong answers or unanswered questions scored 0 points. Assessments were scored by two evaluators (blinded to group allocation) different from the two instructors. We used the kappa statistic to analyze the agreement between evaluators (Altman 1991).

Secondary endpoints were the actual increase in the score between the different times of each assessment (final value – initial value) and the score gain (Δ), defined as the ratio between the actual increase in the score and the potential increase:

$$\Delta_{\text{Test}} = 100 \cdot [(\text{final value} - \text{initial value}) / (3 - \text{initial value})]$$

$$\Delta_{\text{Video}} = 100 \cdot [(\text{final value} - \text{initial value}) / (4 - \text{initial value})]$$

Finally, the number of students who improved their T2 scores from T1 was counted; we considered a percentage of over 50% to be relevant.

Statistics

Statistical calculations were performed using the SPSS 13.0 package (SPSS Inc., Chicago IL, USA). Qualitative variables were expressed as numbers (percentage) and were compared using the chi-square test. Fisher’s exact test was used for a 2×2 frequency table with a number of observed frequencies less than 5. Quantitative variables were not normally distributed. Data were expressed as medians and 25th–75th percentiles. Non-parametric testing was used to assess differences between groups (unpaired two-tailed Mann–Whitney U-test) and within groups (Wilcoxon signed rank sum test for related samples). Statistical significance was defined as $p < 0.05$.

Results

Sixty-eight medical students completed the study, 34 in each group. Two participants in Group I were excluded because they did not complete T2 assessments (Figure 1). Previous BLS training experience was similar in both groups (Table 1).

Table 2 summarizes the results of the assessments. The kappa index of inter-evaluator agreement was higher than 0.8 for all assessments. We found no significant differences between the groups in any of the recorded scores. In both

Table 1. Personal characteristics and pre-study training experience in cardiopulmonary resuscitation (CPR) of the students. Data are shown as (a) number of students (% in each group) or as (b) median (25th–75th percentiles).

| | Group I (n = 34) | Group II (n = 34) | <i>p</i> |
|---|---------------------|----------------------|----------|
| Females/males ^(a) | 27/7 (79.4/20.6) | 26/8 (76.5/23.5) | 0.77 |
| Age ^(b) | 20 (20–21) | 20 (20–21) | 0.65 |
| Students who had received previous CPR training | | | |
| • lectures plus hands-on ^(a) | 8 (23.5) | 11 (32.4) | 0.42 |
| • hands-on ^(a) | 1 (2.9) | 2 (5.9) | 1 |
| Students who had performed CPR on a patient or manikin ^(a) | 10 (29.4) | 13 (38.2) | 0.44 |
| • Number of CPRs performed ^(b) | 0 (0–0.75) | 0 (0–1) | 0.43 |

Table 2. Scores obtained at the assessments before (T1) and after (T2) teaching the BLS algorithms. Values are expressed as medians and 25th–75th percentiles. (#) $p < 0.001$ versus T1 within the group.

| Assessment | Group | T1 | T2 |
|-----------------|----------------------|-----------|-------------------------|
| Test | Group I | 1.5 (1–2) | 2 (2–3) [#] |
| | Group II | 1 (1–2) | 2 (2–3) [#] |
| | p (between groups) | 0.71 | 0.75 |
| Video recording | Group I | 1 (0–1) | 2 (2–3) [#] |
| | Group II | 1 (0–1) | 2 (1.75–3) [#] |
| | p (between groups) | 0.72 | 0.43 |

Table 3. Actual increase in the score and gain of score (Δ) at the assessments; T1: score obtained before teaching the BLS algorithms; T2: score obtained after teaching the BLS algorithms. Data are expressed as median and 25th–75th percentiles.

| Score | Group I ($n = 34$) | Group II ($n = 34$) | p |
|--------------------------------|----------------------|-----------------------|------|
| Test T2 – Test T1 | 1 (0–1) | 1 (0–1.25) | 0.52 |
| Δ (Test T2 – Test T1) | 50 (0–58.33) | 50 (0–83.33) | 0.31 |
| Video T2 – Video T1 | 1.5 (1–2) | 1 (1–2) | 0.25 |
| Δ (Video T2 – Video T1) | 50 (25–66.66) | 50 (0–83.33) | 0.33 |

Table 4. Number of students who achieved better scores (positive ranks), worse scores (negative ranks) or the same scores (tie ranks) at the assessments (Wilcoxon rank test). Values are expressed as number of cases (percentage within the group). (a) $p = 0.012$ within group.

| Score differences | Ranks | Group I ($n = 34$) | Group II ($n = 34$) | p |
|---------------------|----------|------------------------|-----------------------|------|
| Test T2 – Test T1 | Positive | 19 (55.9) ^a | 20 (58.8) | 0.8 |
| | Negative | 2 (5.9) | 4 (11.8) | 0.39 |
| | Tie | 13 (38.2) | 10 (29.4) | 0.44 |
| Video T2 – Video T1 | Positive | 29 (85.3) ^a | 28 (82.3) | 0.74 |
| | Negative | 2 (5.9) | 0 (0.0) | 0.15 |
| | Tie | 3 (8.8) | 6 (17.6) | 0.28 |

groups, T2 scores improved significantly both on the test and on the video (p values < 0.001).

The analysis of the actual increases of scores and the gains of scores on the test and on the video did not reveal any significant differences between the groups (Table 3).

We found no differences between the groups in the numbers of students who improved their T2 scores for either assessment (Table 4). In both groups, the percentages of students who improved their scores were relevant. In Group I, the number of students who improved their scores on the video was significantly higher than those who improved their scores on the test.

Discussion

In this study, the new instructional method based on the discussion of simulated case scenarios of non-traumatic cardiac arrest did not show any advantages versus the traditional method based on the multimedia presentation for the evaluated fields of knowledge. Our findings are consistent

with two previous publications of our group in the fields of internship and residency and continuous medical education in which case discussion method did not demonstrate to be superior (Carrero et al. 2007, 2008). Randomized controlled studies have shown similar results in undergraduate medical education (Koles et al. 2005; Wahlgren et al. 2006).

Case discussion method and problem based learning are not the same. Reliance on teacher-directed versus self-directed learning remains the most obvious difference between the case method and problem based learning (Tärnvik 2007). In the present study we applied the case discussion method because we considered it more suitable to our scheduled teaching program. Moreover due to a more teacher-dependent approach to learning, the case based discussion method is less susceptible to group dysfunction, a well recognized problem inherent in problem based learning (Tärnvik 2007).

Case-based teaching designs are widely variable and we cannot rule out that the case-based discussion method could have obtained better results if it had been applied to a different context or if other parameters had been analysed. In this regard, several papers show the efficacy of problem based scenarios for improving non-technical skills and clinical judgment (Yee et al. 2005; Tiwari et al. 2006).

The important aspect of our study lies in the fact that both methods can be used to achieve equally good results and therefore these two methods are equally valid instructional alternatives. Cost-benefit balance of teaching methods, such as those used in this study, is difficult to assess since depends on the already established organization and available infrastructure of the different centers. In our case, both methods were equally affordable in our organization.

The high rates of agreement between evaluators found in our study are a measure of validation and reproducibility for the assessments used, which could lead to their more widespread of them. Unlike multiple-choice tests, designing the assessments with open answers allowed us to better understand the construction of the response by the student and the analysis of cognitive skills beyond simple recognition.

There were common factors to the two teaching methods used in this study that could explain our equally good results: both methods showed the BLS sequence of action in a realistic setting, either filmed or simulated, which approximated teaching to reality. Moreover, both, video recordings and simulation are instructional tools that particularly motivate students, as they consider these tools to be very effective (Lee et al. 2006; Morgan et al. 2006).

A higher percentage of students in both groups improved their error-pinpointing scores to a higher extent than their right-decision making scores. Perhaps, the video recording used for assessment could have had a visual retention effect on the students, helping them to self-redress their mistakes. It is also possible that we may have put stronger emphasis in teaching BLS maneuvers, chest compressions and ventilation, than in the lateral safety position or the management of the airway, which is a common bias in such courses (Durak et al. 2006).

Interestingly, we found that most students did not reach the maximum score at any of the post-instructional assessments. Our results reflect the difficulty in achieving BLS teaching

objectives with one workshop only. This fact has already been reported in the literature (Madden 2006; Reeder et al. 2006) and reinforces the recommendation that BLS education should be continuous (Chamberlain et al. 2002; Chamberlain & Hazinski 2003; Madden 2006).

Our study presents several drawbacks. This study deals with a narrow topic (only cognitive skills on BLS algorithm) in a limited setting (one 60 min session, with 70 students). Institutional limitations in both, available timetable and resources, restricted the study design to that carried out. A design including the effects of the teaching methods on both type of skills, cognitive and practical, would have given a more approximate assessment of the student's performance in the clinical setting. The lack of differences in cognitive skills with both teaching methods used does not exclude the possibility of differences in practical skills.

It is possible that there were differences between both methods that we could not detect because we aimed at a difference in scoring of at least 25%. We would have had to include many more students to detect smaller differences but it was not possible since the sample size was determined by the number of students doing the course in our Medical School. Including students of several years would have increased the variability of the sample, making the interpretation of the results difficult. Sample size is difficult to increase in medical education research since student numbers are determined by the scheduled teaching program.

The assessment method was basic, with simple questions and limited score range. This simplicity might cause a bias affecting validity and prevents us from generalizing our results. The simplicity of the questions could have conditioned the fact that no differences were detected between the groups but we do not consider it a main factor since the assessment method gave low pre-study scores and detected scoring differences in both groups after BLS algorithm instruction. Universal BLS algorithms are intentionally simple and easy to facilitate learning and application in the majority of circumstances of cardiac arrest (Nolan 2005). Simple questions did not imply simple responses; in our design the students' response reflected the students' capacity to integrate and apply the cognitive skills. We did not compare global scores of the whole assessment because we believe that different fields scoring are not comparable.

One of the difficulties in instructional research studies such as ours is harmonizing the curriculum across the analysed groups but the absence of differences between groups in their BLS training background is an indicator of homogeneity for our student sample. However, since it is impossible to have a double-blind design, we believe it is extremely difficult to exclude the professor effect on the results or a better student's performance due to observation.

One differential aspect between the two instructional methods used was the nature of the students' participation during the instruction—passive (multimedia) as opposed to active (case based discussion). However, in the light of our results, we cannot rule out that some factors may have contributed to the similarity between the methods. One factor could have been the 15-minute question-and-answer session in the Group I although we think that such a short time could

have only minimal influence in the ability to cognitive skills and improve learning of students with such a little previous experience in BLS.

Other limitations of the study were logistics that impeded to assess retention of learning over time and the clinical competence on cognitive skills in clinical practice: important goals of the teaching efforts.

The applicability of our methodology is to be confirmed in further studies in order to compare our results and contrast them in different populations of students. In addition, the potential correlation of our assessments to multiple-choice question tests and evaluation of skills needs to be defined. It would be equally interesting to compare new methods for assessing BLS knowledge, extending the analysis of different cognitive abilities and determining the effect of several instructional methodologies. Long-term studies are required to establish whether the improvement in decision making and error pinpointing enhances, in turn, the students' competence in BLS.

Instructional research in BLS should be based on evidence. In this regard, this study provides a methodological design that can be used to objectively assess the efficacy of instruction in the analysed fields of knowledge. Nevertheless, we have to be humble when interpreting our results considering the flaws of this study.

In conclusion, teaching universal BLS algorithms using either an instructional videotape plus Power-Point presentation or case-based discussion equally improves the level of cognitive BLS skills among medical students.

Acknowledgements

The authors thank Erika Sierra for her statistical advice and Dr J Fontanals and Dr J Tercero for their collaboration.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Notes on contributors

ENRIQUE CARRERO, MD is anaesthesiologist and certified ERC ALS instructor.

CARMEN GOMAR, MD, PhD is Professor of Anaesthesiology and Director of the Skills Laboratory of The Medical School of the University of Barcelona.

WILMA PENZO, MD, PhD is Professor of Psychology and Expert in Medical Education.

NEUS FABREGAS, MD, PhD is Professor of Anaesthesiology and certified ERC ALS instructor.

RICARD VALERO, MD, PhD is Professor of Anaesthesiology, certified ERC ALS instructor and Director of the Course of Simulation for Organ Donor Coordinators.

GERARD SANCHEZ-ETAYO is anaesthesiologist and certified ERC ALS instructor.

References

- Altman DG. 1991. Practical statistics for medical research. New York: Chapman & Hall/CRC.

Baskett PJF, Nolan JP, Handley A, Soar J, Biarent D, Richmond S. 2005. European Resuscitation Council Guidelines for Resuscitation 2005: Section 9. Principles of training in resuscitation. *Resuscitation* 67S1:S181–S189.

Carrero E, Gomar C, Penzo W, Rull M. 2007. Comparison between lecture-based approach and case/problem-based learning discussion for teaching pre-anaesthetic assessment. *Eur J Anaesthesiol* 24:1008–1015.

Carrero EJ, Gomar C, Fábregas N, Penzo W, Castillo J, Villalonga A. 2008. Clase magistral versus aprendizaje basado en caso/problema para la enseñanza del embolismo aéreo en formación médica continuada. *Rev Esp Anesthesiol Reanim* 55:202–209.

Chamberlain D, Hazinski MF. 2003. European Resuscitation Council, American Heart Association, Heart and Stroke Foundation of Canada, Resuscitation Council of Southern Africa, Australia and New Zealand resuscitation council, Consejo Latino-Americano de resucitación. Education in resuscitation: An ILCOR symposium: Utstein Abbey: Stavanger, Norway: June 22–24, 2001. *Circulation* 108:2575–2594.

Chamberlain D, Smith A, Woollard M, Colquhoun M, Handley AJ, Leaves S, Kern KB. 2002. Trials of teaching methods in basic life support (3): Comparison of simulated CPR performance after first training and at 6 months, with a note on the value of re-training. *Resuscitation* 53:179–187.

Durak HI, Certug A, Caliskan A, van Dalen J. 2006. Basic life support skills training in a first year medical curriculum: Six years' experience with two cognitive-constructivist designs. *Med Teach* 28:e49–58.

Handley AJ, Koster R, Monsieurs K, Perkins GD, Davies S, Bossaert L. 2005. European Resuscitation Council Guidelines for Resuscitation 2005. Section 2. Adult basic life support and use of automated external defibrillators. *Resuscitation* 67S1:S7–S23.

Hoke RS, Handley AJ. 2006. A reference basic life support provider course for Europe. *Resuscitation* 69:413–419.

Jordan T, Bradley P. 2000. A survey of basic life support training in various undergraduate health care professions. *Resuscitation* 47:321–323.

Koles P, Nelson S, Stolfi A, Parmelee D, Destephen D. 2005. Active learning in a Year 2 pathology curriculum. *Med Educ* 39:1045–1055.

Lee MT, Jacobs JL, Kamin CS. 2006. Video-enhanced problem-based learning to teach clinical skills. *Med Educ* 40:473–474.

Madden C. 2006. Undergraduate nursing students' acquisition and retention of CPR knowledge and skills. *Nurse Educ Today* 26:218–227.

Morgan PJ, Cleave-Hogg D, Desousa S, Lam-McCulloch J. 2006. Applying theory to practice in undergraduate education using high fidelity simulation. *Med Teach* 28:e10–15.

Nolan J. 2005. European Resuscitation Council guidelines for resuscitation 2005. Section 1. Introduction. *Resuscitation* 67S1:S3–S6.

Philips PS, Nolan JP. 2001. Training in basic and advanced life support in UK medical schools: Questionnaire survey. *Br Med J* 323:22–23.

Reder S, Cummings P, Quan L. 2006. Comparison of three instructional methods for teaching cardiopulmonary resuscitation and use of an automatic external defibrillator to high school students. *Resuscitation* 69:443–453.

Tärnvik A. 2007. Revival of the case method: A way to retain student-centred learning in a post-PBL era. *Med Teach* 29:e32–e36.

Tiwari A, Lai P, So M, Yuen K. 2006. A comparison of the effects of problem-based learning and lecturing on the development of students' critical thinking. *Med Educ* 40:547–554.

Wahlgren CF, Edelbring S, Fors U, Hindbeck H, Stähle M. 2006. Evaluation of an interactive case simulation system in dermatology and venereology for medical students. *BMC Med Educ* 14:6–40.

Yee B, Naik VN, Joo HS, Savoldelli GL, Chung DY, Houston PL, Karatzoglou BJ, Hamstra SJ. 2005. Nontechnical skills in anesthesia crisis management with repeated exposure to simulation-based education. *Anesthesiology* 103:241–248.

Appendix 1. Assessments: questions, fields of knowledge measured, correct answers and scoring criteria

Test

Questions

- (1) On the BLS sequence, which maneuvers should be performed for an adult victim lying on the floor onto his back, unresponsive when you check him, who is breathing normally. You are alone and nobody arrives after your shout for help.
- (2) On the BLS sequence, which maneuvers should be performed for an adult victim lying on the floor onto his back, unresponsive when you check him. You are alone and nobody arrives after your shout for help.
- (3) On the BLS sequence, which maneuvers should be performed for an adult victim lying on the floor onto his back, unresponsive when you check him, who is making agonal gasps. You are alone and nobody arrives after your shout for help and you call 112.

Field measured

- (1) Decision based on BLS sequence.

Correct answers

- (1) Turn him into the recovery position
- (2) Open the airway
- (3) Chest compressions

Score system. One point for each correct answer. Score range: 0–3

Video recording

Question. You are about to see a video recording about the adult BLS sequence with one rescuer. The BLS sequence seen in the footage may contain maneuvers performed incorrectly or not. Please indicate which wrong actions, if any, you can identify.

BLS Step:

Safety
Check responsiveness
Shout for help
Open airway
Check breathing

Video:

Victim and rescuer safe
Shake and shout: unresponsive
Help!!
Head tilt and chin lift
Keeping the airway open,
look, listen, and feel:
not breathing

Call 112

Chest compressions:

Call 112

Heel of left hand located on victim's upper abdomen, fingers of both hands interlocked, rescuer's position vertical, arms straight, pressure down 4–5 cm, release pressure without losing contact between rescuer hands and victim.
Chest compression rate
50 times per minute.

Compression to ventilation ratio

5:2

Stop to recheck the victim

After 5 compressions and 2 breaths

Field measured

- (2) Detection of errors on BLS performance

Correct answers (errors correctly detected)

- (1) Heel of left hand located on victim's upper abdomen. (ERC guidelines: in the center of the victim chest).
- (2) Chest compression rate 50 times a minute. (ERC guidelines: rate of approximately 100 min^{-1}).

- (3) Compression to ventilation ratio 5:2. (ERC guidelines: ratio of 30:2).
- (4) Stop to recheck the victim after 5 compressions and 2 breaths. (ERC guidelines: only if he starts breathing normally).

Scoring system. One point for each correct answer. Score range: 0–4.

9.2. Artículo 2.

Carrero E, Gomar C, Penzo W, Rull M. Comparison between lecture-based approach and case/problem-based learning discussion for teaching pre-anaesthetic assessment. *Eur J Anaesthesiol* 2007; 24:1008-15.

Original Article

Comparison between lecture-based approach and case/problem-based learning discussion for teaching pre-anaesthetic assessment

E. Carrero*, C. Gomar*, W. Penzo†, M. Rull‡

University of Barcelona, Hospital Clínic, Departments of *Anaesthesiology, †Clinical Psychology, Barcelona;
‡Hospital Universitari de Tarragona Joan XXIII, Department of Anaesthesiology, Tarragona, Spain

Summary

Background and objective: The case/problem-based learning discussion method was recently introduced into the theory-based training program for residents run by the Catalan Society of Anaesthesiology. This study was designed to assess and compare its effectiveness with that of the lecture-based approach for teaching pre-anaesthetic assessment, applying an objective tool for knowledge evaluation before and after teaching. **Methods:** A prospective randomized study of two consecutive year groups of first year anaesthesiology residents was conducted. Twenty-nine residents attended a lecture, and 25, a case/problem-based learning discussion session. Their knowledge of pre-anaesthetic assessment was assessed before and after the teaching session with tests on four different clinical cases measuring six fields: (1) 'recognizing clinical data with anaesthetic implications'; (2) 'reasoning clinical data with anaesthetic implications'; (3) 'ASA class'; (4) 'Mallampati class'; (5) 'choice of anaesthetic technique'; (6) 'reasoning choice of anaesthetic technique'. **Results:** Before the teaching session, the lecture group scored significantly higher on field 1 ($P = 0.006$). Both teaching methods improved scores on fields 1, 2 and 4. The case/problem group also improved on fields 3 and 6. After the teaching session, the field 1 score was still significantly higher in the lecture group ($P = 0.005$), and the field 3 score was significantly higher in the case/problem group ($P = 0.044$). **Conclusions:** The effectiveness of lecture and case/problem-based learning discussion differed little in terms of improving participants' immediate knowledge of 'pre-anaesthetic assessment'.

Keywords: EDUCATION; RESEARCH DESIGN; PROSPECTIVE STUDIES; RANDOMIZED CONTROLLED TRIALS; METHODS; EDUCATIONAL MEASUREMENT, knowledge; ANAESTHESIA; INTERNSHIP AND RESIDENCY; TEACHING, lectures, problem-based learning.

Introduction

Pre-anaesthetic assessment is a subject that is common to all training programs for residents in Anaesthesiology. It is a complex process in which information must be sought, compiled and analysed.

As the topic involves more than the mere acquisition of theoretical knowledge, it provides an ideal context for the assessment of the effectiveness of new teaching methods.

The methodological aspect of teaching anaesthesiology is an under-explored research field. The design of assessment methods that are able to measure and compare teaching effectiveness in complex areas such as 'pre-anaesthetic assessment' is especially difficult. Traditionally, a lecture-based approach has been used to transmit theoretical

Correspondence to: Enrique Carrero, Department of Anaesthesiology, Hospital Clínic, University of Barcelona, Villarroel 170, 08036, Barcelona, Spain.
E-mail: ecarrero@clinic.ub.es; Tel: +34 93 2275558; Fax: +34 93 2275454

Accepted for publication 11 November 2006 EJA 4097
First published online 30 January 2007

knowledge, though its effectiveness has been questioned [1,2]. In recent years, new teaching techniques have been introduced with the aim of enhancing learning and professional competence, including simulators, online teaching, seminars, video recordings and problem-based learning discussions. There is currently no consensus on the best teaching method.

Several types of problem-based learning have emerged [3], for the most part, in response to the need to adapt this approach to learning programs. Case/problem-based learning discussion [4] is a method that is easy to implement and readily accepted by students without the need for increased educational resources; indeed, the American Society of Anesthesiologists (ASA) has included case/problem-based learning discussion in its continuous medical education courses since 1991 [5]. However, no studies have been published on the use of case/problem-based learning discussion to teach subjects in resident training programs, for example, the topic of 'pre-anaesthetic assessment'.

This study was designed to assess and compare the effectiveness of the case/problem-based learning discussion method and the traditional lecture method in the teaching of the subject 'pre-anaesthetic assessment', using an objective tool for knowledge evaluation before and after teaching.

Methods

The study population included first year anaesthesiology residents in Catalonia from 2 consecutive years. Participation in the study was voluntary, anonymous and had no effect on academic assessment. Residents who did not complete all the stages of the study were excluded. All residents included in the final analysis (Fig. 1) provided written informed consent. The study took place over 2 days of 2 consecutive years during the 'pre-anaesthetic assessment' class within the theoretical training program for first year anaesthesiology residents run by the Catalan Society of Anaesthesiology. A schematic representation of the study is shown in Figure 2.

The teaching contents of the class (Table 1) were previously defined and agreed on by the teaching committee of the program. The two participating teachers had teaching experience in this topic and met before the class to agree on the procedures to apply.

The teaching session lasted 60 min in both groups. The lecture group was allowed 50 min plus additional 10 min for questions, whereas the case/problem group had 60 min for discussion in class. The teacher who taught the case/problem group was an expert in this teaching technique. The lecture

comprised a narrative account with overhead projection, without the participation of students except for the final 10 min for questions. The case/problem discussed in the case/problem group was a patient with high perioperative risk owing to the medical history, with concomitant medication that needed to be readjusted for surgery and difficult airway management; the patient was first scheduled for elective urological surgery, and months later, required emergency surgery following trauma shock.

Four tests were designed, based on four similar clinical cases which, in pairs, met the teaching contents (Appendix 1). The clinical cases simulated four patients scheduled for surgery. The tests were validated 4 months before the study by the two faculty members, three second year anaesthesiology residents, one third year resident and another fourth year resident, and were scored by the same two raters who would later evaluate participants in the study. The raters confirmed that the tests were comprehensible, had the same level of difficulty and could be completed in 20 min. This validation was used to exclude the data or questions in the tests which caused dissent among the raters, and thus to produce four final tests. On the first day, it was randomly decided, with the residents from the first year of the study, which pair of tests would be taken by each group before the teaching session (tests 1 and 2 = pre-tests) and which ones after the teaching session (tests 3 and 4 = post-tests). The order was repeated for the residents from the second year of the study. The maximum time allowed to complete the pre- and post-tests was 20 min. The pre- and post-tests were completed immediately before and after the teaching session (Fig. 2). Students were not allowed to make comments related to the tests until the post-tests were over.

The variables measured in the tests were six knowledge fields in the topic 'pre-anaesthesia assessment': field 1, 'recognizing clinical data with anaesthetic implications'; field 2, 'reasoning clinical data with anaesthetic implications'; field 3, 'ASA class': ASA physical status classification; field 4, 'Mallampati Class': Mallampati classification of the oropharynx to predict difficult tracheal intubation; field 5, 'choice of anaesthetic technique'; field 6, 'reasoning choice of anaesthetic technique'.

Each field had a fixed number of items for which the total number of correct answers was counted (Appendix 1). Each correct item scored 1 point, while wrong or unanswered items scored 0 points. We calculated the score obtained for each of the fields in the pre- and post-tests, and the difference in score between the post- and pre-tests. We then established whether there were differences in score

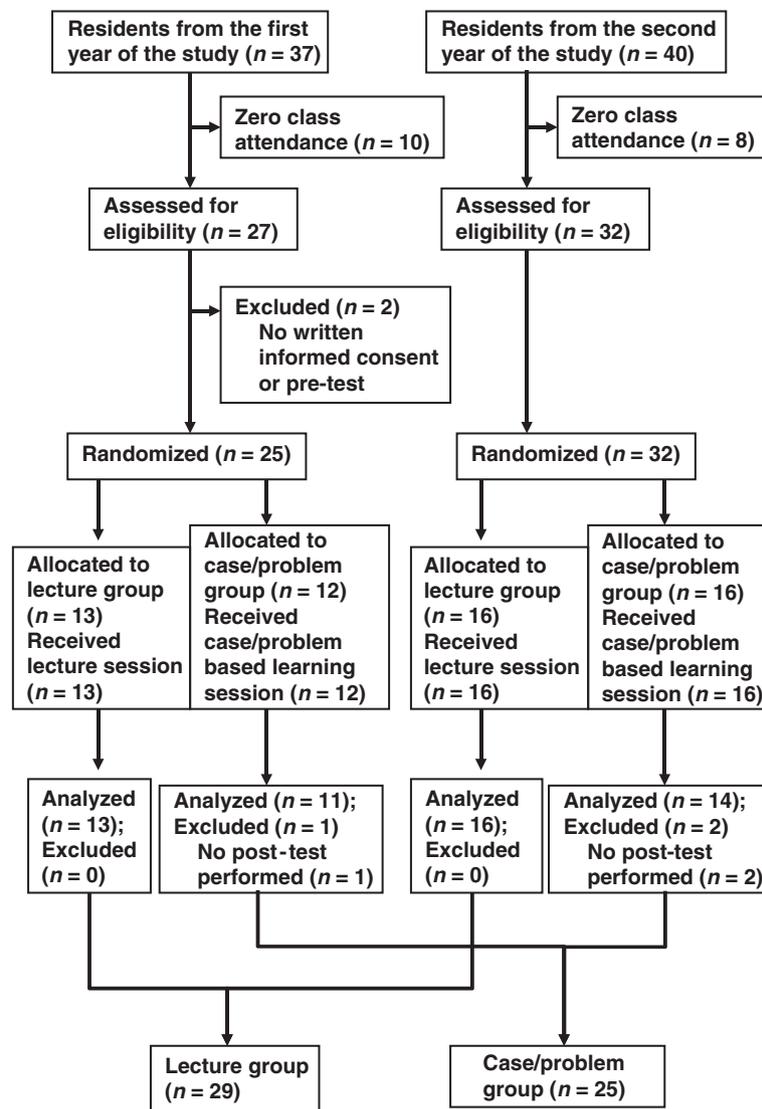


Figure 1.

Flow chart showing participants' progress in the study. Pre-test: evaluation before the teaching session; post-test: evaluation after the teaching session.

between and within groups, and whether differences in the results could be caused by the year of the study considered. All tests were scored by two raters (blinded to group assignment) other than the teachers. Individual items on which no agreement was reached were omitted.

Statistics

Closed envelopes with a number between 1 and 25, for the residents from the first year of the study, and between 26 and 57, for the residents from the second year of the study, were randomly allocated immediately after the pre-test (Fig. 1). Residents were randomized by a computer-generated procedure (two blocking restriction) to receive either a

lecture session (lecture group, 13 residents from the first year of the study and 16 residents from the second year of the study) or a case/problem-based learning discussion session (case/problem group, 12 residents from the first year of the study and 16 residents from the second year of the study). Residents and teachers were not blinded to the educational method applied.

The final sample size was determined by the number of first year anaesthesiology residents in both groups who attended the 'pre-anaesthetic assessment' class, who agreed to participate in the study and who completed the pre- and post-tests. Statistical calculations were performed using the SSPS 10.0 package (SSPS Inc., Chicago, IL, USA). All variables were classified numerically and

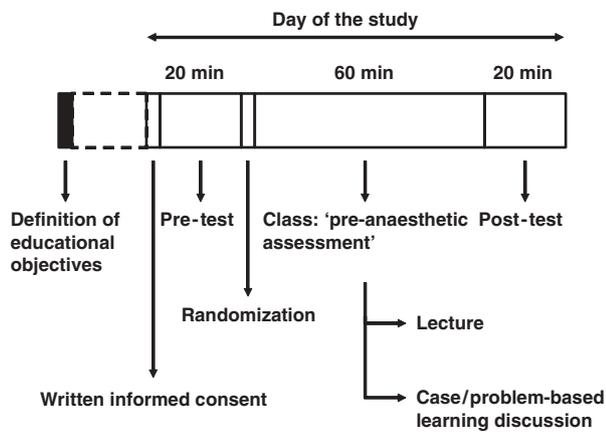


Figure 2. Outline of the methodology followed in the study. Pre-test: evaluation before the teaching session; post-test: evaluation after the teaching session.

Table 1. Teaching contents of the class 'pre-anaesthetic assessment'.

| |
|--|
| Perioperative risk evaluation and measures to reduce morbidity and mortality |
| Analysis of the information obtained from the clinical history |
| Physical examination and complementary studies |
| Assessment of the difficulties encountered in airway management |
| Mallampati classification [6] of the oropharynx to predict difficult tracheal intubation |
| ASA physical status classification [7] |
| Patient optimization before surgery |
| Anaesthetic strategy |
| Patient information and informed consent |
| Pre-medication |
| Analgesia and postoperative care |

interpreted as quantitative variables. Variables were not normally distributed. Data were expressed as median and 25th–75th percentiles. Non-parametric tests were used to assess differences between groups (unpaired two-tailed *U*-test) and within group (Wilcoxon signed rank sum test or McNemar test for dichotomous variables). Statistical significance was set at $P < 0.05$.

Results

The study included 54 residents, 29 in the lecture group and 25 in the case/problem group (Fig. 1). Results of pre- and post-tests are summarized in Table 2. The scores obtained on the pre-tests were similar in the groups for all measured fields except for field 1, which was significantly higher in the lecture group ($P = 0.006$).

After the teaching session, field 1, field 2 and 4 scores improved significantly in the lecture group ($P = 0.001$, $P = 0.001$ and $P = 0.000$, respectively). In the case/problem group, significant improvements were found in field 1 ($P = 0.001$), field 2 ($P = 0.000$), field 3 ($P = 0.042$), field 4 ($P = 0.001$) and field 6 ($P = 0.022$) scores.

The field 1 post-test score was significantly higher in the lecture group ($P = 0.005$), while the field 3 post-test score was significantly higher in the case/problem group ($P = 0.044$). We found no significant differences between the groups for the other fields when comparing post-test scores. Analysis of the difference between pre- and post-test scores revealed no significant differences between the two groups for any of the fields studied.

Table 2. Scores obtained before the teaching session (pre-test) and after the teaching session (post-test) according to fields of knowledge measured.

| Field | Test | Lecture group ($n = 29$) | Case/problem group ($n = 25$) |
|---|------|----------------------------|---------------------------------|
| Field 1 'Recognizing clinical data with anaesthetic implications' | Pre | 9 (8–10)* | 8 (7–9)* |
| | Post | 10 (9.5–11)*# | 10 (8–10)*# |
| Field 2 'Reasoning clinical data with anaesthetic implications' | Pre | 5 (4–6) | 4 (4–6) |
| | Post | 7 (6–9)# | 8 (7–9)# |
| Field 3 'ASA class' | Pre | 1 (0–1.5) | 1 (0–1) |
| | Post | 1 (0–1)* | 1 (1–2)*# |
| Field 4 'Mallampati class' | Pre | 0 (0–1) | 0 (0–1) |
| | Post | 1 (1–1)# | 1 (1–1)# |
| Field 5 'Choice of anaesthetic technique' | Pre | 1 (1–1) | 1 (1–1) |
| | Post | 1 (1–1) | 1 (1–1) |
| Field 6 'Reasoning choice of anaesthetic technique' | Pre | 1 (1–2) | 1 (1–2) |
| | Post | 2 (1–2) | 2 (1–2)# |

Data are expressed as medians and 25th–75th percentiles. **U*-test, $P < 0.05$ between groups. #Wilcoxon signed rank sum test, $P < 0.05$ within group. (n): number of cases analysed.

Discussion

The study shows that both lecture and case/problem-based learning discussion obtain satisfactory results regarding the immediate acquisition of knowledge of 'pre-anaesthetic assessment'. The results demonstrate that case/problem-based learning discussion is a suitable teaching method for this topic and that, unlike other forms of problem-based learning [8], it does not require many organizational changes.

Despite the interest shown in problem-based learning, there is no consistent evidence that this teaching method is superior to the lecture in increasing practitioners' knowledge [8,9]. In our study, case/problem-based learning discussion produced better results only for the acquisition of knowledge about the ASA risk scale. Some authors have reported better reasoning ability with problem-based learning methodology [10,11]. Although our results do not support this conclusion, we did observe that case/problem-based learning discussion improved the two fields related to clinical reasoning, whereas the traditional lecture improved only one. The limitations of our study do not allow to generalize our observation to other topics of the teaching program and there are no data in the literature to recommend either of the two teaching methods for specific topics. Our study tried to implement a tool to assess the results of teaching methods, therefore, the conditions of the study were better controlled by limiting the analysis to only one specific topic. According to our results, we can only conclude that either method can be used to teach a 'pre-anaesthetic assessment' topic; the choice may depend on the preferences of the faculty or the residents themselves [12].

Problem-based learning is now being introduced into anaesthesiology training programs, especially in undergraduate [13] and continuous medical education [5]; however, for resident training, only the applicability of the teaching method and the degree of satisfaction among students have been addressed [14] without exploring knowledge acquisition. As the good acceptance of problem-based learning has already been established [5,15], therefore, we did not assess it in our study.

The prior evaluation of any teaching method is essential for assessing its effectiveness and, if carried out sufficiently in advance, also allows curriculum designers to adjust the teaching contents and improve the learning objectives for each group. In our study, the evaluations were performed immediately before and after the teaching session, so as to make sure that the students did not receive any information on 'pre-anaesthetic assessment' other

than that taught in class. Despite immediate knowledge acquisition being only a part of the complex professional learning process, it offers the possibility to attribute the results obtained to the teaching method applied without external influences. To our knowledge, there are no randomized controlled studies of 'pre-anaesthetic assessment' that include the evaluation of knowledge before and after a teaching session. Therefore, we are unable to compare our results.

As recognizing, reasoning, memorizing and selection are the different processes in the integration and application of knowledge, they were scored separately in our study. We think that the present study introduces a readily applicable and reproducible evaluation model that allows an objective evaluation of the students' knowledge.

A small sample size is a flaw typical of educational studies and one that is difficult to solve. In our study, the sample size and, therefore, its statistical power, was determined by the number of residents attending the class. Fifty-four first year anaesthesiology residents could be considered small from the perspective of statistical power, but relevant in terms of representation specially compared with data in the literature.

No assessment methods have been described that measure all facets of clinical competence, that is, knowledge, skills and attitudes. Indeed, we did not evaluate skills or attitudes and, therefore, we do not know whether the teaching session helped enhance the participants' clinical competence. Moreover, the potential for improving the communication skills of each method, essential in the pre-anaesthetic professional performance assessment, was not evaluated. Rodrigues de Oliveira demonstrated the applicability and effectiveness of a checklist based on answers to performance items to evaluate residents' competence in 'pre-anaesthetic assessment' [16]. To be able to conclude whether a resident is capable of performing 'pre-anaesthetic assessment', we must have instruments to measure how he/she performs in a real situation (performance assessment *in vivo*) [17].

Other methodological drawbacks in studies of this kind, and also in ours, include the difficulty of studying homogeneous adult groups, the impossibility of conducting blind studies [18], the possible improvement in performance, students and teachers, when subjects are aware of being observed [19] and the difficulty of evaluating long-term knowledge retention.

In our study, the lecture group had scored higher in the field 'recognizing clinical data with anaesthetic implications', in the initial evaluation. Though participants' knowledge improved in the

same proportion in both groups, this means that the two groups were not homogeneous, emphasizing the fact that in adults education, a previous evaluation of knowledge is required to objectively assess the effectiveness of any teaching method. Although the teaching session improved participants' knowledge in this study, we were surprised by the previous high scores in both groups. The finding suggests that the teaching contents of 'pre-anaesthetic assessment' are, for the most part, already known by medical postgraduates; indeed, the contents covered are basic medical issues [20] and, therefore, the teaching objective here should not be to increase residents' knowledge, but rather to encourage them to apply it to the perioperative medicine approach.

Education requires a great deal of planning, time and resources. The design of studies of this kind presents considerable methodological difficulties, but research into the effectiveness of educational interventions in training anaesthesiology residents must be based on objective evidence. The uses of fields of knowledge with pre- and post-teaching session evaluations can be a useful tool to quantify and compare the effectiveness of different teaching methods in the area of 'pre-anaesthetic assessment'. In this group of first year anaesthesiology residents, the effectiveness of lecture and case/problem-based learning discussion differed little in terms of improving participants' immediate knowledge of the topic of 'pre-anaesthetic assessment'. More studies are needed to test our results for other topics in anaesthesiology, and to determine the effect of different teaching methodologies on long-term retention of knowledge, skills, attitudes and clinical competence. An appropriate methodology for this research must be developed.

Acknowledgements

The authors thank the Catalan Society of Anaesthesiology, specially Dr G Fita, for facilitating the organization and structure to develop this study, Erika Sierra for her statistical advice and Dr J Castillo, Dr N Fábregas and Dr R Valero for their assistance in reviewing the manuscript.

References

1. Willenkin RL. Lectures in anesthesia training [editorial]. *Anesth Analg* 1992; 74: 1–2.
2. Landers DF, Becker GL, Newland MC, Peters KR. Lecture practices in United States anesthesiology residencies. *Anesth Analg* 1992; 74: 112–115.
3. Barrows H, Tambly R, eds. *Problem Based Learning*. New York, USA: Springer, 1980.
4. Barrows HS. A taxonomy of problem-based learning methods. *Med Educ* 1986; 20: 481–486.
5. Liu PL. 1996 Problem-based learning discussions program progress to new pinnacles. *ASA Newsletter* [serial online] 1996 July; 60(7): [3 screens]. Available from: URL: http://www.asahq.org/Newsletters/1996/07_96/Article10.htm.
6. Mallampati SR. Clinical signs to predict difficult tracheal intubation (hypothesis). *Can Anaesth Soc J* 1983; 30: 316–317.
7. Wolters U, Wolf T, Stützer H, Schröder T. ASA classification and perioperative variables as predictors of postoperative outcome. *Br J Anaesth* 1996; 77: 217–222.
8. Colliver JA. Educational theory and medical education practice: a cautionary note for medical school faculty. *Acad Med* 2002; 77: 1217–1220.
9. Smits PB, de Buissonje CD, Verbeek JH, van Dijk FJ, Metz JC, ten Cate OJ. Problem-based learning versus lecture-based learning in postgraduate medical education. *Scand J Work Environ Health* 2003; 29: 280–287.
10. Karaalp A, Akici A, Kocabasoglu YE, Oktay S. What do graduates think about a two-week rational pharmacotherapy course in the fifth year of medical education? *Med Teach* 2003; 25: 515–521.
11. Lohse B, Nitzke S, Ney DM. Introducing a problem-based unit into a lifespan nutrition class using a randomized design produces equivocal outcomes. *J Am Diet Assoc* 2003; 103: 1020–1025.
12. Dugdale A. Traditional medical education and the new path – they are not mutually exclusive. *Med Educ* 2001; 35: 304.
13. Stehr SN, Muller M, Frank MD *et al.* Teaching methods in anesthesia and intensive care medicine. The new legislation and its possibilities for the specialty. *Anaesthesist* 2005; 54: 385–393.
14. Rodrigues de Oliveira G, Schonhorst L. Problem-based learning implementation in an intensive course of anaesthesiology: a preliminary report on residents' cognitive performance and perceptions of the educational environment. *Med Teach* 2005; 27: 382–384.
15. Chang CH, Yang CY, See LC, Lui PW. High satisfaction with problem-based learning for anaesthesia. *Chang Gung Med J* 2004; 27: 654–662.
16. Rodrigues de Oliveira G, Schonhorst L. The development and application of an instrument for assessing resident competence during preanesthesia consultation. *Anesth Analg* 2004; 99: 62–69.
17. Whitcomb ME. More on competency-based education. *Acad Med* 2004; 79: 493–494.
18. Norman GR, Schmidt HG. Effectiveness of problem-based learning curricula: theory, practice and paper darts. *Med Educ* 2000; 34: 721–728.
19. Draper SW. *The Hawthorne effect and other expectancy effects: a note* [www document] 2005 March [15 screens]. Available from: URL: <http://www.psy.gla.ac.uk/~steve/hawth.html>.
20. Feeley TW. The anesthesiologist as perioperative physician, new challenges, new roles. In: Schwartz J, Matjasko J, Otto Ch, eds. *ASA Refresher Courses in Anesthesiology*. Orlando, USA: Lippincott-Raven Publishers, 1998; 224: 1–7.

Appendix 1

Pre-tests and post-tests: questions, fields of knowledge measured, correct answers and score criteria

Pre-tests

Test 1: An 84-yr-old patient with a history of poorly controlled hypertension and congestive heart failure. One month previously, he had an explorative laparotomy for acute abdomen, suffering a myocardial infarction in the immediate postoperative period. Emergency intervention for colectomy.

Test 2: A 60-yr-old female patient, 155 cm, 120 kg, allergic to latex. History of bronchial asthma, requiring treatment with inhaled bronchodilators and prednisone 10 mg day⁻¹ for several years. Physical exploration: limited mouth opening, only tongue and hard palate visible. Baseline blood gas analysis: PaO₂ 56 mmHg. Scheduled for hiatus hernia surgery.

Post-tests

Test 3: Full-term pregnancy, requiring Caesarean section due to delay in labour. Family history of breast neoplasm. Smoker, one cigarette pack per day for past 10 yr. Appendectomy 4 yr previously. Mallampati reveals hard palate and a small part of soft palate without pillars or uvula.

Test 4: A 50-yr-old patient. Aortic valve replacement 2 yr previously, with a good functional result; since then treated with oral anticoagulants. Emergency surgery required for tibia and fibula open fractures. Physical examination revealed cutaneous-mucous pallor, tachycardia 130 beats min⁻¹ and systemic blood pressure (BP) 75/50 mmHg. Mallampati Class I.

Question 1

'Underline the items in the text which in your view have implications for planning patient anaesthesia'.

Field measured: Field 1, 'recognizing clinical data with anaesthetic implications'.

Correct answers (test 1): 'age 84', 'hypertension', 'heart failure', 'myocardial infarction', 'colectomy'.

Correct answers (test 2): 'allergy to latex', 'asthma', 'prednisone', 'limited mouth opening', 'PaO₂ 56 mmHg', 'hiatus hernia'.

Correct answers (test 3): 'full-term pregnancy', 'Caesarean section', 'smoker', 'Mallampati score'.

Correct answers (test 4): 'aortic valve replacement', 'oral anticoagulants', 'tibia and fibula open fractures', 'pallor', 'tachycardia', 'systemic BP 75/50 mmHg', 'Mallampati Class I'.

Score system for pre- and post-tests: 1 point for item underlined correctly. Score range: 0–11.

Question 2

'For each item that you underlined, explain why you think it is relevant for anaesthesia'.

Field measured: Field 2, 'reasoning clinical data with anaesthetic implications'.

Correct answers: a correct reasoning for each one of the items correctly underlined.

Score system for pre- and post-tests: 1 point for each item correctly underlined and a correct explanation.

Score range: 0–11.

Question 3

'Evaluate the risk for the patient according to the ASA classification'.

Field 3, 'ASA class'.

Correct answer (test 1): 'ASA Class IV'.

Correct answer (test 2): 'ASA Class III'.

Correct answer (test 3): 'ASA Class II'.

Correct answer (test 4): 'ASA Class III'.

Score system for pre- and post-tests: 1 point for each correct item. Score range: 0–2.

Question 4 (only for tests 2 and 3):

'What is the patient's Mallampati classification?'

Field 4, 'Mallampati class'.

Correct answer (test 2): 'Mallampati Class IV'.

Correct answer (test 3): 'Mallampati Class III'.

Score system for pre- and post-tests: 1 point for each correct item. Score range: 0–1.

Question 5 (only for tests 1 and 3):

'What anaesthetic technique (general or local/regional) would you choose in this case?'

Field 5 'Anaesthetic technique choice'.

Correct answer (test 1): 'local/regional anaesthesia'.

Correct answer (test 3): 'local/regional anaesthesia'.

Score system for pre- and post-tests: 1 point for each correct item. Score range: 0–1.

Question 6 (for tests 1 and 3):

'Explain your choice of anaesthetic technique'.

Question 6 (for tests 2 and 4):

'Why do you think the best technique for this patient is general anaesthesia?'

Field 6, 'Reasoning choice of anaesthetic technique'.

Correct answers: A correct reasoning to indicate local/regional anaesthesia (for test 1 and 3) or general anaesthesia (for tests 2 and 4).

Score system for pre- and post-tests: 1 point for each item correctly reasoned. Score range: 0–2.

9.3. Artículo 3.

Carrero E, Gomar C, Fábregas N, Penzo W, Castillo J, Villalonga A. Clase magistral versus aprendizaje basado en caso/problema para la enseñanza del embolismo aéreo en formación médica continuada. Rev Esp Anesthesiol Reanim 2008; 55: 202-9.

Clase magistral *versus* aprendizaje basado en caso/problema para la enseñanza del embolismo aéreo en formación médica continuada

E. J. Carrero^{a,b}, C. Gomar^{a,b}, N. Fábregas^a, W. Penzo^c, J. Castillo^{b,d}, A. Villalonga^{b,e}

^aServicio de Anestesiología y Reanimación, Hospital Clínic, Universidad de Barcelona. ^bFundación Europea para la Enseñanza en Anestesiología.

^cDepartamento de Psicología Clínica, Universidad de Barcelona. ^dServicio de Anestesiología y Reanimación, Hospital de l'Esperança, Barcelona.

^eServicio de Anestesiología y Reanimación, Hospital Universitari de Girona Dr Josep Trueta, Girona.

Resumen

OBJETIVO: La eficacia de la formación médica continuada (FMC) en Anestesiología es un área poco explorada. Este estudio compara la eficacia de la clase magistral con el “aprendizaje basado en caso/problema” (ABCP) para la enseñanza del embolismo aéreo en FMC.

MATERIAL Y MÉTODOS: Estudio prospectivo y randomizado. Se incluyeron 52 anestesiólogos con experiencia clínica que participaron en un curso de FMC. Veintisiete anestesiólogos recibieron una clase magistral y 25 una sesión de ABCP sobre “embolismo aéreo en anestesia”. Los objetivos docentes se definieron con antelación y fueron los mismos para ambos grupos. El conocimiento de los participantes se evaluó antes y después de la enseñanza con tests basados en dos casos clínicos diferentes que exploraban los mismos campos: “factores de riesgo y síntomas”, “diagnóstico”, “monitorización” y “tratamiento”.

RESULTADOS: No se encontraron diferencias significativas entre los grupos antes de la enseñanza ni después de la misma en ninguno de los campos. Después de la enseñanza, el grupo clase magistral mejoró sus puntuaciones en “monitorización” ($p = 0,03$) y “tratamiento” ($p = 0,001$); el grupo caso/problema mejoró sus puntuaciones en “tratamiento” ($p = 0,003$).

CONCLUSIONES: No hubo diferencias en el conocimiento de los participantes entre los dos grupos de aprendizaje; las mejorías alcanzadas fueron mínimas. La metodología del estudio permitió una valoración objetiva de los conocimientos adquiridos.

Palabras clave:

Educación médica continuada. Clases magistrales. Aprendizaje basado en problemas. Conocimiento. Estudios controlados randomizados. Anestesiología. Embolismo aéreo.

Problem/case-based learning compared to lectures for acquiring knowledge of air embolism in continuing medical education

Summary

OBJECTIVE: The efficacy of continuing medical education in anesthesiology has been examined very little. This study compared the efficacy of a lecture on air embolism to that of a class that used a problem/case-based learning approach.

MATERIAL AND METHODS: Prospective, randomized study enrolling 52 experienced anesthesiologists participating in a professional development course. Twenty-six anesthesiologists attended a lecture on air embolism in anesthesia and 25 attended a problem-based class. The objectives were the same for both groups and had been defined previously. The participants' knowledge was evaluated before and after the instruction with tests based on 2 cases dealing with the same knowledge areas: risk factors and symptoms, diagnosis, monitoring, and treatment.

RESULTS: No significant between-group differences were found for any of the knowledge areas before or after the classes. After instruction, participants who listened to the lecture improved their scores for knowledge of monitoring ($P = .03$) and treatment ($P = .001$). Participants in the problem-based learning group also improved their scores for knowledge of treatment ($P = .003$).

CONCLUSIONS: No between-group differences in participants' knowledge outcomes were detected; improvements were minimal. The study design allowed the knowledge acquired to be evaluated objectively.

Key words:

Continuing medical education. Lecture classes. Problem-based learning. Knowledge. Randomized controlled trial. Anesthesiology. Air embolism.

Correspondencia:

Dr Enrique J Carrero
Servicio de Anestesiología y Reanimación
Hospital Clínic.
Villarroel 170
08036 Barcelona.
E-mail: ecarrero@clinic.ub.es; 24566ecc@comb.es

Aceptado para su publicación en enero de 2008.

Introducción

Son escasas las investigaciones que se han realizado en formación médica continuada (FMC) en Anestesiología¹⁻³.

Desde 1986, la Fundación Europea para la Enseñanza en Anestesiología (FEEA) viene desarrollando

un programa de FMC de amplia aceptación⁴. Cada curso organizado por el centro FEEA en Cataluña, incluye 30 horas lectivas que se imparten en 2,5 días. El método docente más frecuentemente utilizado es la clase magistral.

Las diversas modalidades del método “aprendizaje basado en problemas” han surgido, en su mayoría, por la necesidad de adaptar este método docente a los programas de enseñanza. El aprendizaje basado en casos/problema⁵ (ABCP) es un ejemplo. La sociedad Americana de Anestesiólogos viene incluyendo el ABCP en sus cursos de FMC desde 1991; el método ha sido bien aceptado por los participantes y su implantación no ha precisado costes adicionales⁶. En el centro FEEA en Cataluña, hemos introducido recientemente el ABCP para determinados contenidos del programa docente. No hemos encontrado en la literatura estudios prospectivos que comparen la eficacia de la clase magistral y el ABCP en FMC en Anestesiología. En otros campos de FMC los resultados de estudios comparativos similares han sido contradictorios⁷.

El embolismo aéreo en anestesia es una complicación potencialmente fatal donde el anestesiólogo juega un papel crucial en reducir la morbi-mortalidad gracias a un diagnóstico y tratamiento precoces⁸⁻¹⁰; por este motivo, éste es un tema esencial en los programas de FMC en Anestesiología. Tradicionalmente se ha enseñado en el formato clase magistral, pero el método ABCP puede ser también un formato adecuado¹¹.

El objetivo del presente estudio fue comparar la eficacia de la clase magistral y el ABCP en la adquisición de conocimientos sobre embolismo aéreo en anestesia como parte de nuestro programa de FMC.

Material y métodos

Sujetos

Para el cálculo del tamaño muestral, se estimó una eficacia media del 75%; se consideró relevante una diferencia en las puntuaciones del 25%; se estableció un riesgo $\alpha = 0,05$, una potencia estadística $(1-\beta) = 0,8$ y se estimó una pérdida de casos del 10%. El tamaño muestral resultante fue de 29 participantes en cada grupo. La población del estudio fueron 59 anestesiólogos en ejercicio profesional inscritos en el Curso 5 del Centro Cataluña de la FEEA. El programa trató temas relacionados con el sistema nervioso, la anestesia regional y el tratamiento del dolor. La participación en el estudio fue voluntaria y anónima y a todos los participantes se les entregó una hoja de consentimiento informado. Los criterios de inclusión en el estudio fue-

ron: firma del consentimiento para participar en el estudio, asistencia a la clase “embolismo aéreo en anestesia” y cumplimentación del pretest y el post-test. Se excluyeron aquellos participantes que no cumplieron todas las fases del estudio. La Figura 1 muestra el diagrama de randomización y participación (diagrama de flujo) y la Figura 2 las etapas del estudio.

Enseñanza

Los objetivos docentes del tema “embolismo aéreo en anestesia” fueron definidos con detalle y consensuados con anterioridad al estudio por el comité docente del centro. Los dos profesores que impartieron la clase eran miembros del comité. Los objetivos incluyeron: definición de embolismo aéreo, su incidencia, factores de riesgo, fisiopatología, manifestaciones clínicas, diagnóstico, monitorización, pronósti-

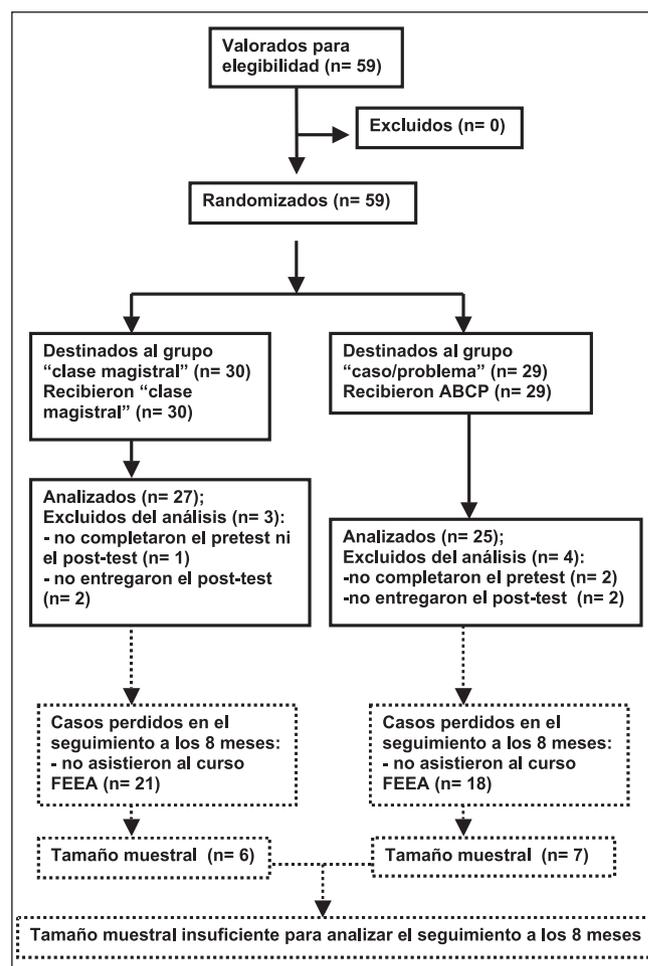


Fig. 1. Diagrama de flujo que muestra la progresión de los participantes en el estudio. ABCP: aprendizaje basado en caso/problema. Pretest: evaluación antes de la enseñanza; post-test: evaluación después de la enseñanza. FEEA: Fundación Europea para la Enseñanza de la Anestesiología.

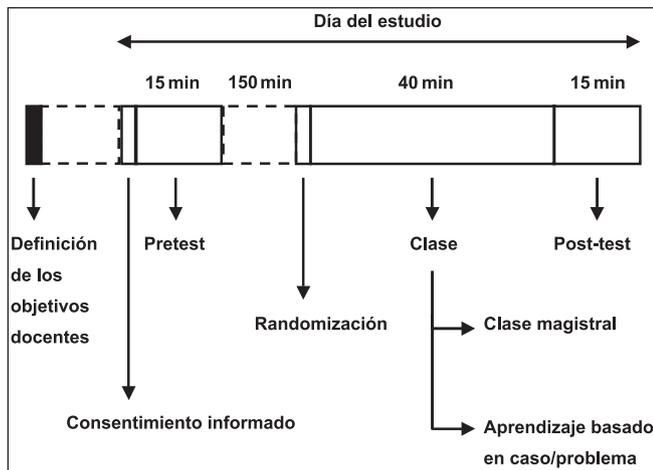


Fig. 2. Esquema de la metodología seguida en el estudio. Pretest: evaluación antes de la enseñanza; post-test: evaluación después de la enseñanza.

co, prevención, tratamiento, y embolismo aéreo paradójico, de acuerdo a la literatura especializada⁸⁻¹¹.

Ambas clases duraron 40 minutos. En el grupo clase magistral la clase incluyó 30 minutos de lección magistral más 10 minutos para preguntas y respuestas, en el grupo caso/problema los 40 minutos fueron de discusión. El material de soporte docente fue una presentación de Power Point en el grupo clase magistral y una pizarra en el grupo caso/problema. El caso clínico discutido en el grupo caso/problema fue un paciente en respiración espontánea que presentó un episodio de embolismo aéreo durante la cirugía de su enfermedad de Parkinson¹². La discusión basada en el caso/problema se presentó como una sesión única y los participantes no recibieron el caso para su preparación por adelantado.

Evaluación

La evaluación se basó en las respuestas a dos tests referidos a dos casos clínicos hipotéticos (Anexo 1) que cumplieran los mismos objetivos docentes. Ambos casos simulaban un embolismo aéreo durante un procedimiento neuroquirúrgico. No se ha descrito un sistema de evaluación y puntuación estandarizado para el embolismo aéreo. El formato y contenido de los tests fueron similares al subalgoritmo del embolismo aéreo validado por Williamson et al.⁸ para el manejo de situaciones críticas en anestesia. Previamente al estudio, se realizó una prueba piloto en un grupo de anesestesiólogos que no participaron como alumnos en el curso; los tests fueron puntuados por los dos mismos evaluadores que posteriormente puntuaron a los participantes del estudio. Se confirmó que los tests eran comprensibles, del mismo grado de dificultad y fáciles de completar en 15 minutos. Esta prueba piloto sirvió para eliminar de los tests los ítems o preguntas que

causaban discrepancia entre los profesores. El mismo día del estudio se decidió al azar cuál de los dos tests se tenía que cumplimentar antes de la enseñanza (pre-test) y cuál después de la enseñanza (post-test). Antes de la clase, no se proporcionó a los participantes ningún material docente del tema “embolismo aéreo”, a diferencia del resto de temas del curso. El pretest se cumplimentó 150 minutos antes de la clase y el post-test se realizó inmediatamente después de la clase. El tiempo máximo permitido para la cumplimentación de cada test fue de 15 minutos. Los participantes sólo pudieron realizar preguntas relacionadas con los casos clínicos de los tests cuando finalizó el post-test.

Las variables medidas en los tests fueron cuatro campos de conocimiento sobre el tema “embolismo aéreo en anestesia”: 1) Factores de riesgo y síntomas; 2) Diagnóstico; 3) Monitorización; y 4) Tratamiento. Cada campo incluía un número determinado de ítems sobre los que se contabilizó el número total de respuestas correctas (Tabla 1). Los ítems contestados correctamente puntuaron 1 punto y los ítems no contestados o contestados incorrectamente puntuaron 0 puntos. Se calculó la puntuación obtenida para cada uno de los campos en el pretest y en el post-test. También se calculó la diferencia en la puntuación “post-test menos pretest”. Por último, se analizó si existían diferencias entre los grupos e intragrupo. Los tests fueron puntuados por dos evaluadores diferentes a los dos profesores, ciegos a la asignación de los grupos. Empleamos el estadístico kappa¹³ para analizar la concordancia entre evaluadores.

Para evaluar la retención de conocimientos, se intentó repetir el pretest 8 meses más tarde, coincidiendo con el inicio de otro curso en el que se preveía la inscripción de gran parte de los participantes. Sin embargo, sólo 13 participantes se inscribieron en el nuevo curso (6 pertenecientes al grupo clase magistral y 8 pertenecientes al grupo caso/problema) por lo que los datos resultaron insuficientes para su análisis.

Análisis estadístico

Inmediatamente después del pretest, se distribuyeron sobres cerrados a los participantes con los números del 1 al 59. Los participantes quedaron randomizados por el ordenador (restringido a dos bloques) para recibir una clase magistral (grupo clase magistral, 30 participantes) o una sesión de ABCP (grupo caso/problema, 29 participantes). Los participantes y los profesores no fueron ciegos al método de enseñanza aplicado.

Los cálculos estadísticos se realizaron usando el paquete estadístico SSPS 10.0 (SSPS Inc, Chicago IL, USA). Las variables cualitativas se expresaron en

TABLA 1

Preguntas, respuestas correctas, criterios y rangos de puntuación correspondientes a los cuatro campos del conocimiento medidos en los tests

| | | |
|---|--|--|
| Campo 1: Factores de riesgo y síntomas. <i>“Por favor marque los ítems que, en su opinión, tienen relación con la complicación expuesta y justifique porqué”</i> | | |
| Respuestas correctas | Pretest | “Posición sentada”; “coincidiendo con un episodio de sangrado venoso”; “de forma brusca”; “descenso en el ETCO ₂ a 6 mmHg”; “taquicardia”; “hipotensión”. |
| | Post-test | “Abordaje posterior con cabeza incorporada”; “de forma súbita”; “marcada depresión del segmento ST”; “descenso del ETCO ₂ a 10 mmHg en un minuto hasta llegar finalmente a 0 mmHg”; “descenso de la presión arterial a valores indetectables”; “disminución de las cifras de SpO ₂ hasta perderse el registro” |
| Puntuación | 1 punto por cada ítem correcto más 1 punto por cada ítem correcto con razonamiento correcto | |
| Rango de puntos | 0-12 | |
| Campo 2: Diagnóstico. <i>“Establezca el diagnóstico diferencial. Señale su diagnóstico de sospecha”.</i> | | |
| Respuestas correctas | Pretest | “Embolismo aéreo” |
| | Post-test | “Embolismo aéreo” |
| Puntuación | 1 punto por incluir “embolismo aéreo” en el diagnóstico diferencial más 1 punto por señalar “embolismo aéreo” como diagnóstico de sospecha | |
| Rango de puntos | 0-2 | |
| Campo 3: Monitorización. <i>“A su juicio, qué monitorización sería útil para confirmar su diagnóstico de sospecha”.</i> | | |
| Respuestas correctas | Pretest | “Capnografía”; “ecocardiografía”; “fonendoscopio o doppler precordial”; “catéter aurícula derecha” |
| | Post-test | “Capnografía”; “ecocardiografía”; “fonendoscopio o doppler precordial”; “catéter aurícula derecha” |
| Puntuación | 1 punto por cada ítem correcto | |
| Rango de puntos | 0-4 | |
| Campo 4: Tratamiento. <i>“Escriba qué medidas terapéuticas adoptaría ante la complicación descrita en el caso clínico”.</i> | | |
| Respuestas correctas | Pretest | “Avisar, tapar entrada de aire, hemostasia”; “parar N ₂ O, incrementar FiO ₂ ”; “soporte hemodinámico”; “aspirar aire por catéter venoso central”; “cambiar posición del paciente” |
| | Post-test | “Avisar, tapar entrada de aire, hemostasia”; “parar N ₂ O, incrementar FiO ₂ ”; “soporte hemodinámico”; “aspirar aire por catéter venoso central”; “cambiar posición del paciente” |
| Puntuación | 1 punto por cada ítem correcto | |
| Rango de puntos | 0-5 | |

números absolutos (proporciones) y se compararon usando el test Chi-cuadrado. El test exacto de Fisher se empleó para las tablas de frecuencias 2x2 cuando los valores de las frecuencias esperadas fueron inferiores a 5. Las variables cuantitativas no siguieron una distribución normal. Los datos se expresaron como medianas y percentiles 25-75. Se utilizaron pruebas no paramétricas para analizar las diferencias entre grupos (test de la U de Mann-Whitney para muestras independientes) e intragrupo (test de Wilcoxon para muestras apareadas). Definimos la significación estadística como $p < 0,05$.

Resultados

Excluimos 7 participantes de los 59 iniciales por no cumplimentar o entregar alguno de los tests. Cincuenta y dos participantes se incluyeron finalmente en el estudio, 27 en el grupo clase magistral y 25 en el grupo caso/problema (Figura 1). Los resultados del pretest y post-test están resumidos en la Tabla 2. El índice kappa de concordancia entre evaluadores fue superior a 0,8 para todas las evaluaciones.

No se encontraron diferencias estadísticamente significativas entre los grupos en el pretest ni en el post-test para ninguno de los campos.

La comparación intra grupo de las puntuaciones del pretest y post-test mostró incrementos significativos en el grupo clase magistral en “monitorización” (test de Wilcoxon, $p = 0,03$) y “tratamiento” ($p = 0,001$). El grupo caso/problema sólo obtuvo incrementos significativos en “tratamiento” ($p = 0,003$). No se encontraron diferencias significativas entre los dos grupos al comparar la diferencia en las puntuaciones “post-test menos pretest” para ninguno de los campos.

El análisis por separado de las puntuaciones “marcar los ítems” y “justificar porqué” del campo “factores de riesgo y síntomas” no mostró diferencias significativas intragrupo ni entre los grupos en las puntuaciones pretest y post-test.

El análisis desglosado de la puntuación “diagnóstico” mostró que la inclusión del ítem “embolismo aéreo” en el diagnóstico diferencial mejoró significativamente en el post-test en el grupo caso/problema (test de Wilcoxon para muestras apareadas, $p = 0,046$). El porcentaje de participantes que incluyeron el ítem “embolismo aéreo” en el diagnóstico diferencial en el post-test fue significativamente mayor en el grupo caso/problema ($25/25 = 100\%$) que en el grupo clase magistral ($21/27 = 77,78\%$) (test exacto de Fisher, $p = 0,023$).

El análisis desglosado de la puntuación “monitorización” mostró que la respuesta “fonendoscopio o

TABLA 2

Puntuaciones en el pretest y post-test para los campos del conocimiento evaluados

| Campo | Test | Grupo clase magistral (n = 27) | Grupo caso/problema (n = 25) |
|-------------------------------|-----------|--------------------------------|------------------------------|
| Factores de riesgo y síntomas | pretest | 8 (6-11) | 8 (6-11) |
| | post-test | 9 (6-10) | 8 (6-9,5) |
| Diagnóstico | pretest | 2 (1-2) | 2 (1-2) |
| | post-test | 2 (1-2) | 2 (1-2) |
| Monitorización | pretest | 1 (1-2) | 1 (1-2,5) |
| | post-test | 2 (1-3)# | 2 (1-2) |
| Tratamiento | pretest | 2 (1-4) | 3 (2-4) |
| | post-test | 4 (3-5)# | 4 (3,5-5)# |

Los datos estan expresados en medianas y percentiles 25-75. (#) Test de Wilcoxon para muestras apareadas, $p < 0,05$ intragrupo

doppler precordial” mejoró significativamente en el post-test tanto en el grupo clase magistral (test de Wilcoxon para muestras apareadas, $p = 0,002$) como en el grupo caso/problema ($p = 0,035$). La respuesta “catéter aurícula derecha” disminuyó significativamente en el post-test en el grupo caso/problema ($p = 0,034$). No se encontraron diferencias significativas entre los grupos en el porcentaje de participantes que indicó cada una de las monitorizaciones en ninguno de los tests.

El análisis desglosado de la puntuación “tratamiento” mostró que la respuesta “parar N₂O, incrementar FiO₂” mejoró significativamente en el post-test tanto en el grupo clase magistral (test de Wilcoxon para muestras apareadas, $p = 0,000$) como en el grupo caso/problema ($p = 0,001$); la respuesta “aspirar aire por catéter venoso central” mejoró significativamente en el grupo clase magistral ($p = 0,035$); y las respuestas “avisar, tapar entrada de aire, hemostasia”, ($p = 0,021$), “soporte hemodinámico” ($p = 0,025$), y “cambiar posición del paciente”, ($p = 0,02$) mejoraron significativamente en el grupo caso/problema. Se encontraron diferencias significativas entre los dos grupos en el post-test en las respuestas “cambiar posición del paciente” (grupo clase magistral: $16/27 = 59,26\%$; grupo caso/problema: $22/25 = 88\%$), (test chi-cuadrado, $p = 0,02$) y “aspirar aire por catéter venoso central” (grupo clase magistral: $22/27 = 81,48\%$; grupo caso/problema: $8/25 = 32\%$), (test chi-cuadrado, $p = 0,000$).

Discusión

El hallazgo principal del estudio fue que los conocimientos adquiridos por los alumnos fueron similares con la clase magistral y el método de ABCP; y aunque entre el pretest y el post-test existieron diferencias

estadísticamente significativas en dos campos, éstas fueron menores. No hemos encontrado estudios controlados y randomizados en FMC en Anestesiología que nos permitan comparar nuestros resultados.

En nuestro estudio, el alto grado de acuerdo entre los dos evaluadores es una medida de validación y reproducibilidad de los tests aplicados, que podría favorecer la difusión de su uso. Hasta ahora no se ha descrito un método válido y universal para la evaluación de la eficacia de la FMC¹⁴. En Anestesiología, en particular, no hay referencias sobre cuál es el mejor método de evaluar los conocimientos¹⁵. Los exámenes de respuesta múltiple son el método de evaluación más comúnmente empleado, sobre todo por su medición objetiva de las respuestas. Sin embargo, no valoran la construcción de la respuesta por parte del alumno, su capacidad de aplicar el conocimiento ni la competencia profesional¹⁶. El diseño de los tests en formato de caso clínico aproximó las evaluaciones a la realidad clínica, y la clasificación de las preguntas por campos de conocimiento favoreció el seguimiento de los objetivos docentes.

El diseño de los tests fue simple. Se ha demostrado que la aplicación de algoritmos sencillos, explícitamente estructurados para el embolismo aéreo durante la anestesia favorecen que el diagnóstico sea más precoz y el manejo terapéutico mejor⁸. Cabría preguntarse si debido a el formato simple las preguntas de los tests pudieron, de hecho, detectar diferencias entre los grupos. En nuestro caso pudo influir en la valoración del campo “diagnóstico” que sólo incluía dos ítems, pero creemos que no fue un factor determinante porque el mismo formato detectó diferencias después de la enseñanza en ambos grupos. No calculamos una puntuación global para cada test porque, en nuestra opinión, las puntuaciones de los diferentes campos no eran comparables. Recientemente, hemos aplicado con éxito el mismo formato de evaluación el programa de formación teórica de residentes¹⁷.

El “aprendizaje basado en problemas” ha despertado mucho entusiasmo, sin embargo, no se ha encontrado evidencia de que este método sea más eficaz que la clase magistral en mejorar los conocimientos y habilidades¹⁸; en este sentido, nuestros resultados son coincidentes. Los datos publicados en Anestesiología y en otros ámbitos de las ciencias de la salud sólo evidencian que el “aprendizaje basado en problemas” produce un mayor grado de satisfacción de los alumnos^{2,7,19}. La satisfacción de los participantes con un determinado método de enseñanza podría influir en su eficacia. Nosotros no evaluamos este aspecto y esto constituye una limitación del estudio.

No podemos descartar que la ausencia de diferencias entre los grupos estuviera influenciada por factores distintos a la eficacia de los métodos docentes. El

hecho de que los participantes en ambos grupos fuesen profesionales con experiencia y con un elevado nivel de conocimientos previos es un factor importante capaz de atenuar las posibles diferencias debidas al método docente; de igual forma, estas diferencias pudieron estar atenuadas por el efecto discusión de los diez minutos para preguntas y respuestas del grupo clase magistral.

El pretest mostró que el conocimiento previo de los participantes sobre el tema ya era elevado, lo que disminuyó, sin duda, la posibilidad de detectar grandes mejoras en el post-test. Esta es la diferencia entre la FMC y la formación reglada: en FMC es difícil definir las necesidades de un grupo de profesionales y la formación debe ser adaptada casi de modo individual. Es interesante que después de la enseñanza no se obtuvieran las puntuaciones máximas en los cuatro campos analizados. Los resultados reflejan la dificultad de conseguir los objetivos docentes con una sola clase de 40 minutos y va a favor, por tanto, de mantener un programa de formación continuada que incluya el “embolismo aéreo” dentro de las situaciones críticas en Anestesia. No descartamos la posibilidad de que existan otros métodos de enseñanza más eficaces que la clase magistral o el ABCP para este tema. Los programas de entrenamiento basados en simulaciones de alta fidelidad han demostrado ser efectivos en mejorar las habilidades no técnicas como el trabajo en equipo o la capacidad de decisión²⁰; los participantes también perciben un cambio en su práctica clínica como consecuencia de este tipo de formación^{1,3}. Sin embargo, estos programas conllevan un elevado coste económico, difícil de obtener, porque requieren un entorno físico sofisticado, equipamiento y personal entrenado²¹. Nosotros intentamos mejorar la eficacia de la enseñanza sin cambios organizativos sustanciales.

La investigación en educación médica no es fácil porque se enfrenta con limitaciones metodológicas difíciles de superar y nuestro estudio no fue una excepción. Es posible que realmente existieran diferencias entre los dos métodos de enseñanza que no detectásemos por haber prefijado una diferencia de al menos un 25% entre las puntuaciones. Detectar diferencias más pequeñas implicaría incrementar enormemente el número de participantes (¡275 alumnos por grupo para detectar, por ejemplo, diferencias del 10%!). A diferencia de los estudios clínicos donde es posible incluir más o menos pacientes, en los estudios docentes el tamaño de la muestra está condicionado por el número de alumnos adecuados al plan de formación, que en el caso de la FEEA son 50 participantes por curso. La alternativa de un estudio multicéntrico para incluir un mayor número de alumnos es compleja por la gran variabilidad existente entre los distintos centros de for-

mación. Es muy difícil estandarizar el curriculum de los participantes; en nuestro caso la ausencia de diferencias en el pretest fue un índice de homogeneidad entre los grupos. La randomización se restringió a dos bloques para que el número de casos en cada grupo fuese el máximo y optimizar así el poder estadístico de las mediciones. La variabilidad posible dentro de cada método docente es alta; en nuestro caso, tuvimos que adaptar el método ABCP a las condiciones del programa docente y al propio diseño del estudio. Que los participantes y los profesores conozcan los métodos docentes es una limitación innata a este tipo de estudios²². En nuestro programa de FMC, los participantes estaban acostumbrados al método tradicional de clase magistral mientras el ABCP fue una novedad. Descubrimos si un mayor tiempo de adaptación en el grupo caso/problema hubiera mejorado los resultados como sugieren algunos autores²³. Aunque el número reducido de alumnos en el grupo clase magistral pudo mejorar los resultados, no se ha podido demostrar que un grupo pequeño, como factor independiente, mejore los resultados de la clase magistral en FMC²⁴. Que uno de los grupos mejorase o empeorase más los ítems de un campo sugiere que cada profesor pudo insistir más o menos en dichos aspectos durante su explicación. En estudios de investigación docente como el nuestro, interactúan multitud de variables difíciles de controlar: el contexto donde se realiza la enseñanza, el efecto de los profesores sobre los resultados y la motivación o las expectativas personales de los participantes; todas ellas pudieron interferir en los resultados^{22,25}.

El efecto de los métodos docentes sobre la retención del conocimiento en el tiempo y los cambios en la práctica profesional después de la FMC son aspectos críticos cuando se valora la eficacia. Los estudios de seguimiento de la retención del conocimiento presentan dificultades logísticas y la pérdida de casos es una limitación frecuente en los estudios a medio y largo plazo²⁶. De hecho, en nuestro estudio fuimos incapaces de reclutar suficientes participantes 8 meses después para evaluar la retención de los conocimientos. Nuestro estudio no evaluó habilidades clínicas y los resultados no permiten conocer el impacto de los dos métodos de enseñanza en la competencia clínica.

Las limitaciones mencionadas nos hacen conscientes de que la interpretación de nuestros resultados ha de realizarse con cautela. A pesar de las limitaciones, el presente estudio aporta un diseño metodológico que permite la evaluación objetiva de la eficacia de los métodos de enseñanza de acuerdo a los principios de la FMC basada en la evidencia²⁷. La metodología docente ha de estar supeditada a los criterios de la praxis profesional para que la información resultante sea de calidad. En el campo de la Anestesiología, es

ANEXO 1 Pretest y post-test

Pretest

Mujer de 49 años, peso: 60 Kg, altura: 159 cm; ingesta de alcohol: 20 g día⁻¹; hipertensa en tratamiento con captopril. Diagnosticada por TAC y RMN de meningioma del ángulo ponto-bulbo-cerebeloso; inició tratamiento con corticoides y diuréticos. Se programa para cirugía de la fosa posterior en posición sentada. Mantenimiento anestésico: propofol, fentanilo y vecuronio. Craneotomía sin complicaciones. Durante la exéresis del tumor, coincidiendo con un episodio de sangrado venoso y de forma brusca, se objetivó un descenso en el ETCO₂ a 6 mmHg, taquicardia e hipotensión arterial.

1. "Por favor, marque los ítems que, en su opinión, tienen relación con la complicación expuesta y justifique por qué".
2. "Establezca el diagnóstico diferencial. Señale su diagnóstico de sospecha".
3. "A su juicio, qué monitorización sería útil para confirmar su diagnóstico de sospecha".
4. "Escriba qué medidas terapéuticas adoptaría ante la complicación descrita en el caso clínico".

Post-test

Varón de 55 años, peso: 80 Kg, altura 177 cm; fumador de un paquete de cigarrillos al día desde hace 20 años. Intolerancia gástrica a la aspirina. Colectectomía laparoscópica hace 2 años. Se programa para fijación cervical instrumentalizada por inestabilidad cervical traumática; abordaje posterior con cabeza incorporada. Se cateterizan: vena subclavia derecha, 2 venas periféricas y arteria radial izquierda. El paciente permanece estable las 2 primeras horas de la intervención. Sangrado en 2 horas: 300 mL; líquidos perfundidos en 2 horas: 1L de cristaloides y 0,5 L de coloides. A las 2 horas y 5 minutos del inicio de la intervención se registró de forma súbita: marcada depresión del segmento ST; descenso del ETCO₂ a 10 mmHg en un minuto hasta llegar finalmente a 0 mmHg; descenso de la presión arterial a valores indetectables y disminución de las cifras de SpO₂ hasta perderse el registro.

1. "Por favor, marque los ítems que, en su opinión, tienen relación con la complicación expuesta y justifique por qué".
2. "Establezca el diagnóstico diferencial. Señale su diagnóstico de sospecha".
3. "A su juicio, qué monitorización sería útil para confirmar su diagnóstico de sospecha".
4. "Escriba qué medidas terapéuticas adoptaría ante la complicación descrita en el caso clínico".

muy importante identificar cuáles son las necesidades de aprendizaje de los especialistas y qué factores motivan cambios en su práctica clínica. Sería interesante estudiar la eficacia de la clase magistral y el ABCP en otras áreas de interés de la Anestesiología, en la valoración de las habilidades, aptitudes y competencia profesional. Por último, faltan estudios sobre la influencia de la FMC en Anestesiología en la morbilidad, pronóstico o grado de satisfacción de los pacientes.

En conclusión, la clase magistral y el ABCP fueron igual de eficaces en mejorar los conocimientos inmediatos de los participantes sobre el tema "embolismo aéreo en anestesia". Son precisos más estudios que confirmen y permitan generalizar nuestros resultados. La investigación docente es esencial para mejorar la eficacia de la FMC en Anestesiología.

Agradecimientos

Los autores agradecen a Erika Sierra, Unidad de Bioestadística de la Universidad de Barcelona, su inestimable ayuda en el soporte estadístico del estudio.

BIBLIOGRAFÍA

1. Blum RH, Raemer DB, Carroll JS, Sunder N, Feinstein DM, Cooper JB. Crisis-resource management training for an anaesthesia faculty: a new approach to continuing education. *Med Educ.* 2004;38(1):45-55.
2. Rosenblatt MA. The educational effectiveness of problem-based learning discussions as evaluated by learner-assessed satisfaction and practice change. *J Clin Anesth.* 2004;16(8):596-601.
3. Weller J, Wilson L, Robinson B. Survey of change in practice following simulation-based training in crisis management. *Anaesthesia.* 2003;58(5):471-9.
4. Scherpereel P. Foundation for european education in anaesthesiology (FEEA). *Eur J Anaesthesiol.* 2000;17(2):75-6.
5. Barrows HS. A taxonomy of problem-based learning methods. *Med Educ.* 1986;20(6):481-6.
6. Liu PL. 1996 Problem-based learning discussions program progress to new pinnacles. *ASA newsletter [serie en Internet].* 1996 July; 60 (7): [3 páginas]. Disponible en: URL; http://www.asahq.org/Newsletters/1996/07_96/Article10.htm
7. Smits PB, de Buissonje CD, Verbeek JH, van Dijk FJ, Metz JC, ten Cate OJ. Problem-based learning versus lecture-based learning in postgraduate medical education. *Scand J Work Environ Health.* 2003;29(4):280-7.
8. Williamson JA, Helps SC, Westhorpe RN, Mackay P. Crisis management during anaesthesia: embolism. *Qual Saf Health Care.* 2005;14(3):e17.
9. Domaigne CM. Anaesthesia for neurosurgery in the sitting position: a practical approach. *Anaesth Intensive Care.* 2005;33:323-31.
10. Smith DS, Osborn I. Posterior fossa: anesthetic considerations. En: Cottrell JE, Smith DS, (Eds.). *Anesthesia and Neurosurgery*, 4th ed. St Louis: Mosby; 2001:335-51.
11. Hernández C, Villalonga A. Embolia gaseosa en una paciente intervenida de un tumor de fosa posterior. En: Gomar C, Villalonga A (Eds.). *Casos Clínicos Anestesiología II*, 2ª ed. Barcelona: Masson; 2005:480-8.

12. Suarez S, Ornaque I, Fábregas N, Valero R, Carrero E. Venous air embolism during Parkinson surgery in patients with spontaneous ventilation. *Anesth Analg*. 1999;88(4):793-4.
13. Altman DG. *Practical statistics for medical research*. New York: Chapman & Hall /CRC; 1991.
14. Sibley JC, Sackett DL, Neufeld V, Gerrard B, Rudnick KV, Fraser W. A randomized trial of continuing medical education. *NEJM*. 1982; 306:511-5.
15. Lampe M. Appraisal and reassessment of the specialist in anaesthesia. *Best Pract Res Clin Anaesthesiol*. 2002;16(3):391-400.
16. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA*. 2002;287(2):226-35.
17. Carrero E, Gomar C, Penzo W, Rull M. Comparison between lecture-based approach and case/problem-based learning discussion for teaching the topic 'pre-anaesthetic assessment'. *Eur J Anaesthesiol*. Epub 2007;30:1-8.
18. Smits PB, Verbeek JH, de Buissonjé CD. Problem based learning in continuing medical education: a review of controlled evaluation studies. *BMJ*. 2002;324:153-6.
19. Lucas M, García Guasch R, Moret E, Llasera R, Melero A, Canet J. El aprendizaje basado en problemas aplicado a la asignatura de pregrado de Anestesiología, Reanimación y Terapéutica del Dolor. *Rev Esp Anesthesiol Reanim*. 2006;53:419-25.
20. Dugdale A. Traditional medical education and the new path—they are not mutually exclusive. *Med Educ*. 2001;35(3):304.
21. Yee B, Naik VN, Joo HS, Savoldelli GL, Chung DY, Houston PL, et al. Nontechnical skills in anesthesia crisis management with repeated exposure to simulation-based education. *Anesthesiology*. 2005;103(2):241-8.
22. Kurrek MM, Devitt JH. The cost for construction and operation of a simulation centre. *Can J Anaesth*. 1997;44(11):1191-5.
23. Prideaux D. Researching the outcomes of educational interventions: a matter of design. *BMJ*. 2002;324(7330):126-7.
24. Albanese M. Problem-based learning: why curricula are likely to show little effect on knowledge and clinical skills. *Med Educ*. 2000;34(9):729-38.
25. Davis DA, Thomson MA, Oxman AD, Haynes RB. Evidence for the effectiveness of CME. A review of 50 randomized controlled trials. *JAMA*. 1992;268(9):1111-7.
26. Davis D, O'Brien MA, Freemantle N, Wolf FM, Mazmanian P, Taylor-Vaisey A. Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? *JAMA*. 1999;282:867-74.
27. Herzig S, Linke RM, Marxen B, Borner U, Antepohl W. Long-term follow up of factual knowledge after a single, randomised problem-based learning course. *BMC Med Educ*. 2003;3:3.
28. Wong AK. Curriculum development in anesthesia: basic theoretical principles. *Can J Anesth*. 2006;53:950-60.

10. COMUNICACIONES A REUNIONES CIENTÍFICAS

E Carrero, C Gomar, W Penzo, N Fabregas, A Villalonga J Castillo. Comparison of two teaching methods for continuous medical education in anaesthesia. Comunicación oral y póster en el *22nd Annual Meeting of the European Academy of Anaesthesiology*. Paris (Francia), 21-23 septiembre 2000. Resumen publicado en *Eur J Anaesthesiol* 2000; 17, 754-85.

E Carrero, C Gomar, W Penzo. Diseño de un método de investigación para tres áreas y tres niveles de docencia en Anestesiología. Póster en las 5es Jornades de Docència. Facultad de Medicina. UB. 10-11 Octubre 2000.

E Carrero, C Gomar, W Penzo, N Fábregas, A Villalonga, J Castillo. Comparación de dos métodos de enseñanza en formación continuada en anestesiología. Comunicación oral en el XXV Congreso de la Sociedad Española de Anestesiología, Reanimación y Terapia del Dolor. Tenerife, 6-9 junio 2001. Resumen publicado en *Rev Esp Anesthesiol Reanim* 2001; 48 (Supl 1): 28-9.

E Carrero, C Gomar, W Penzo, M Rull. Case problem based discussion compared with traditional lecturing improves Anesthesiology Residents year1 comprehensive learning in preoperative assessment. Comunicación oral y póster en el *11th International Ottawa Conference on Medical Education*. Barcelona, 6-8 julio 2004. Resumen publicado en *Educ Med* 2004; 7 (3): 88.

Carrero E, Gomar C, Penzo W, Fábregas N, Valero R, Sánchez-Etayo G. Teaching basic life support algorithms by either multimedia presentations or case based discussion equally improves the level of cognitive skills of undergraduate medical students. Póster en el *9th Scientific Congress of the European Resuscitation Council (ERC)*. Ghent, Bélgica. 22-24 mayo 2008.

COMPARISON OF TWO TEACHING METHODS FOR CONTINUOUS MEDICAL EDUCATION IN ANESTHESIA

E. Carrero, C. Gomar, W. Penzo, N. Fabregas, A. Villalonga, J. Castillo.

*FEEA Regional Center of Cataluña,
Department of Anesthesiology and Psychology,
Hospital Clínic, University of Barcelona,
Barcelona, Spain.*



Aim of the study

- To evaluate the effectiveness of two teaching methods of Continuous Medical Education (CME) in Anesthesia.

Methods (I)

- 59 anesthesiologists registered in FEEA, Regional Center of Cataluña, course V (Neurology) were randomly assigned to receive either a didactic (**Group I**) or interactive (**Group II**) CME teaching session on air embolism (AE).

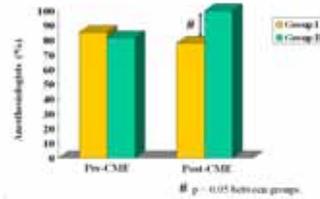
Methods (II)

- Two different case report tests were carried out before and post-CME sessions to assess anesthesiologists' knowledge about AE:
 - A - Case report issues.
 - B - Differential diagnosis.
 - C - Monitoring.
 - D - Treatment.

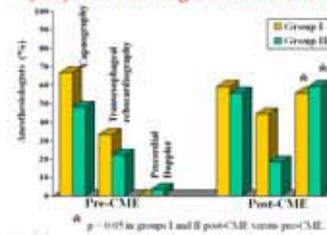
Results (I)

- No difference in pre-teaching knowledge (A,B,C,D) was detected between groups.
- After CME sessions the ability to identify issues related to AE (A) did not change.

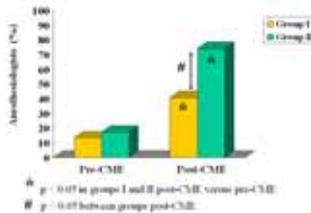
(II): AE in differential diagnosis.



(III): Monitoring considered for AE.



(IV): Algorithm treatment for AE.



CONCLUSIONS

- CME increased knowledge about AE monitoring (precordial doppler) and treatment.
- The interactive teaching method improved anesthesiologists' performance.

References

- JAMA 1992; 268: 1111-7.
- Eur J Anesthesiol 2000; 17: 75-6.

Diseño de un método de investigación para tres áreas y tres niveles de docencia de Anestesiología.

Enrique Carretero, Carmen Gomar, Wilma Penzo
*Departamento de Anestesiología y Psicología,
Hospital Clínic,
Universidad de Barcelona.*

OBJETIVO

- Establecer un método de investigación útil para valorar la eficacia de diferentes sistemas docentes aplicados a la enseñanza de la Anestesiología y Reanimación.

METODOLOGÍA



Grupos de estudio

- Pregrado:** Estudiantes de medicina que reciben clases de Anestesia.
- Postgrado:** MIR de Anestesiología y Reanimación.
- Formación Médica Continuada:** Anestesiólogos con experiencia clínica.

Áreas de docencia

- Valoración preoperatoria.
- Reanimación cardiopulmonar.
- Situación crítica intraoperatoria.

Métodos de enseñanza

- Grupo I: método tradicional**
 - teórico, didáctico o magistral.
- Grupo II: método caso/problema**
 - enseñanza basada en casos/problemas.
 - sesiones interactivas.
 - enseñanza con maniqués.

Valoración de conocimientos

- Casos problema:**
 - similitud real
 - niveles de complejidad
 - criterios críticos y secundarios
- Videos**
 - habilidades
 - detección errores
 - criterio de corrección

↓
**Capacidad discriminativa y de razonamiento.
Conocimientos prácticos.**

Análisis de resultados

- Categorización de las respuestas.**
 - variables directas (cualitativas y cuantitativas)
 - generación de nuevas variables
- Matriz de datos:**
 - programa estadístico SPSS.
- Diferencias intragrupo e intergrupo.**

Limitaciones

- Adaptación a los programas docentes.
- Riesgo de sesgo:**
 - necesidad de participación de un equipo de docentes.
 - contaminación exógena de conocimientos.
- Pérdida de casos en el tiempo.

CONCLUSIONES

- La investigación en docencia resulta fundamental para mejorar la calidad de la enseñanza.

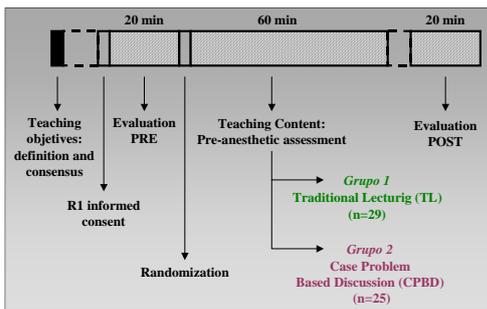
Case problem based discussion (CPBD) compared with traditional lecturing (TL) improves Anesthesiology Residents year 1 comprehensive learning in preoperative assessment.

Objective

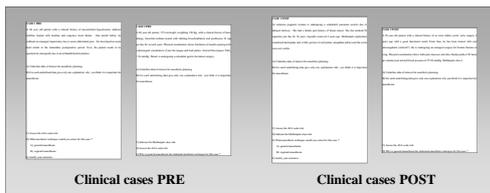
- To compare the effectiveness of two teaching methods for pre-anesthetic skills assessment:
 - Group 1: Traditional Lecturing (TL).
 - Group 2: Case Problem Based Discussion (CPBD).

Methods (I)

- A controlled, prospective, randomized, single blind study.
- **Sample:** Residents of Anesthesiology on the first year (R1) from a geographical region with 6 million inhabitants (Cataluña).
- **Field of the study:** A 60 minutes class of pre-anesthetic assessment.
- **Teaching methods:** Traditional Lecturing (TL) group 1 (n=29) versus Case Problem Based Discussion (CPBD) group 2 (n=25).



Cronology of the study



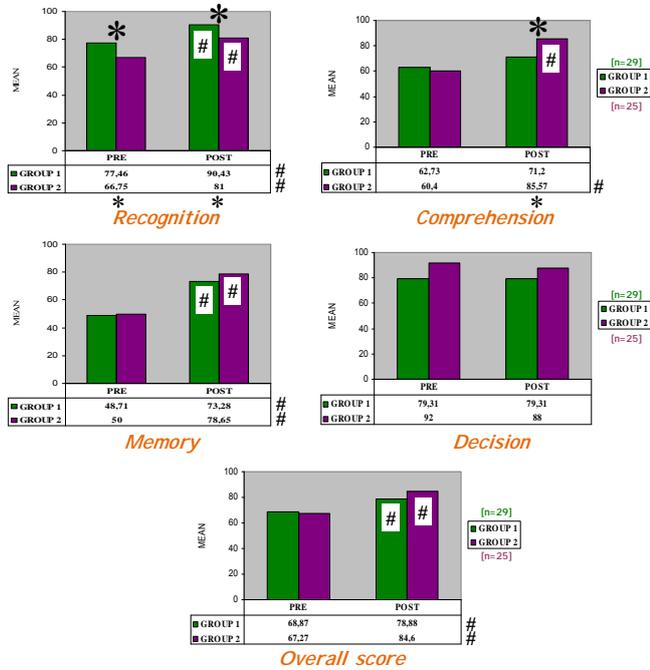
Evaluations



Methods (II)

- **Evaluation:** Two clinical cases PRE and POST (20 minutes per evaluation) with questions related to: *recognition* of data of interest for anesthetic planning; *comprehension*: correct interpretation for recognized data and anesthetic technique indicated; *memory*: physical status (ASA) and difficult intubation (Mallampati) scales of preoperative risk; *decision*: correct selection of the anesthetic technique.
- **Statistics:** SPSS 10.0 for Windows. Qualitative variables: Chi-Square tests. Differences between groups (*): Mann-Whitney U test. Differences within group (#): Wilcoxon Signed Ranks test, Sign test or McNemar (dicotomic variables) test. Level of significance 0.05.

Results



Conclusion

- **CPBD** was more effective than **TL** for pre-anesthetic skills assessment as **CPBD** produced better comprehensive learning.

Teaching BLS algorithms by either multimedia presentations or case based discussion equally improves the level of applied knowledge of undergraduate medical students



Carrero E, Gomar C, Penzo W, Fábregas N, Valero R, Sánchez-Etayo G
DEPARTMENT OF ANAESTHESIOLOGY. HOSPITAL CLINIC.
UNIVERSITY OF BARCELONA



Aim of the study: To compare two methods of teaching BLS, assessing the level of applied knowledge.



Methods: Randomized, prospective study including 70 medical students. BLS algorithms were taught for 60 minutes using either a multimedia presentation (Group I, n = 34) or case based discussion (Group II, n = 34). Assessments included a scenario-based quiz test and an error-pinpointing video, which the students completed before (T1) and after (T2) teaching. Comparisons between both groups were made on scores of the assessments, actual increases in scores (final value - initial value) and score gains (actual increase / potential increase).

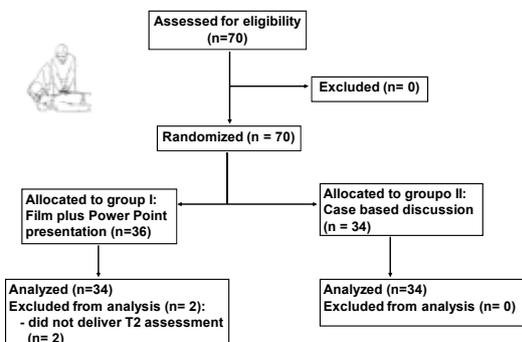


Figure 1. Flow chart showing the medical students' progress in the study.

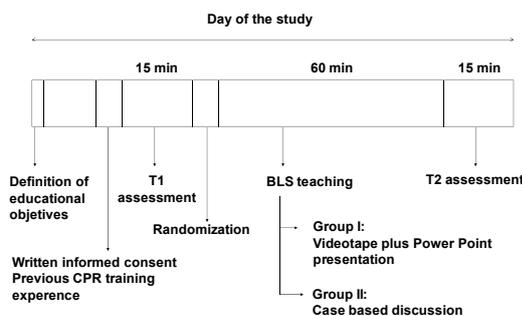


Figure 2. Outline of the methodology followed in the study.

Results:

Table 1. Personal characteristics and pre-study training experience in cardiopulmonary resuscitation (CPR) of the students.

| | Group I (n = 34) | Group II (n = 34) | p |
|---|------------------|-------------------|------|
| Females / males ^(a) | 27/7 (79.4/20.6) | 26/8 (76.5/23.5) | 0.77 |
| Age ^(b) | 20 (20-21) | 20 (20-21) | 0.65 |
| Students who had received previous CPR training | | | |
| - lectures plus hands-on ^(a) | 8 (23.5) | 11 (32.4) | 0.42 |
| - hands-on ^(a) | 1 (2.9) | 2 (5.9) | 1 |
| Students who had performed CPR on a patient or manikin ^(a) | 10 (29.4) | 13 (38.2) | 0.44 |
| - Number of CPRs performed ^(b) | 0 (0-0.75) | 0 (0-1) | 0.43 |

Data are shown as (a) number of students (% in each group) or as (b) median (25th - 75th percentiles).

Table 3. Actual increase in the score and gain of score (Δ) at the assessments;

| Score | Group I (n = 34) | Group II (n = 34) | p |
|--------------------------------|------------------|-------------------|------|
| Test T2 - Test T1 | 1 (0-1) | 1 (0-1.25) | 0.52 |
| Δ (Test T2 - Test T1) | 50 (0-58.33) | 50 (0-83.33) | 0.31 |
| Video T2 - Video T1 | 1.5 (1-2) | 1 (1-2) | 0.25 |
| Δ (Video T2 - Video T1) | 50 (25-66.66) | 50 (0-83.33) | 0.33 |

T1: score obtained before teaching the BLS algorithms; T2: score obtained after teaching the BLS algorithms. Data are expressed as median and 25th - 75th percentiles.

Table 2. Scores obtained at the assessments before (T1) and after (T2) teaching the BLS algorithms.

| Assessment | Group | T1 | T2 |
|-----------------|--------------------|-----------|-------------------------|
| Test | Group I | 1.5 (1-2) | 2 (2-3) [#] |
| | Group II | 1 (1-2) | 2 (2-3) [#] |
| | p (between groups) | 0.71 | 0.75 |
| Video recording | Group I | 1 (0-1) | 2 (2-3) [#] |
| | Group II | 1 (0-1) | 2 (1.75-3) [#] |
| | p (between groups) | 0.72 | 0.43 |

Values are expressed as medians and 25th - 75th percentiles. (#) p < 0.001 versus T1 within the group.

Table 4. Number of students who achieved better scores (positive ranks), worse scores (negative ranks) or the same scores (tie ranks) at the assessments (Wilcoxon rank test).

| Score differences | Ranks | Group I (n = 34) | Group II (n = 34) | p |
|---------------------|----------|------------------------|-------------------|------|
| Test T2 - Test T1 | Positive | 19 (55.9) ^a | 20 (58.8) | 0.8 |
| | Negative | 2 (5.9) | 4 (11.8) | 0.39 |
| | Tie | 13 (38.2) | 10 (29.4) | 0.44 |
| Video T2 - Video T1 | Positive | 29 (85.3) ^a | 28 (82.3) | 0.74 |
| | Negative | 2 (5.9) | 0 (0.0) | 0.15 |
| | Tie | 3 (8.8) | 6 (17.6) | 0.28 |

Values are expressed as number of cases (percentage within the group). (a) p = 0.012 within group.

Conclusion: BLS teaching by either multimedia presentations or case based discussion equally improves the level of applied knowledge among medical students.



