

Elementos para la valoración y uso práctico de los ensayos clínicos. Parte II: Búsqueda, valoración y uso de los resultados

Juan Carlos Villar Centeno*

RESUMEN

Un ensayo clínico aleatorizado (ECA) bien elaborado provee información al personal de salud interesado en obtener información válida sobre el efecto de intervenciones en salud que busca prevenir o tratar un problema de salud. Los usuarios de los ECA deben por tanto conocer las fuentes confiables que les permitan acceder fácil y rápidamente a los ECA que buscan resolver las preguntas surgidas de su propia práctica clínica. Una vez identificado un ECA relevante para el problema, un lector con elementos que le permitan juzgar su validez y aplicabilidad hará un uso más adecuado de ellos. Usando un ECA publicado en MEDUNAB como referencia, este artículo presenta: a) Las razones por las cuales se recurre a ECA para guiar las decisiones clínicas sobre intervenciones en salud, usando como patrón los llamados niveles de evidencia y grados de recomendaciones sobre intervenciones; b) Algunas fuentes de fácil y amplio acceso a la mejor evidencia proveniente de ECA usando Internet, específicamente la biblioteca Cochrane, algunas revistas secundarias y el servicio PubMed; c) Las herramientas para juzgar la calidad de los ECA, en donde se discuten los elementos recomendados por el grupo de trabajo medicina basada en la evidencia y la escala de Jadad; d) Algunas medidas comunmente utilizadas para la estimación del efecto de las intervenciones, tales como las medidas de riesgo absoluto y relativo, y el número necesario a tratar; e) Finalmente, se discuten algunos de los criterios para la aplicación de los ECA en la práctica clínica diaria.

Palabras clave: Ensayos clínicos, metodología.

Introducción

Repasadas las bases de la estructura y la terminología usada en los ensayos clínicos aleatorizados (ECA) en la primera parte de este artículo, nos dedicaremos en esta segunda parte a la discusión de los elementos necesarios para buscar, seleccionar, interpretar y –si es del caso– utilizar la información contenida en este tipo de estudios. Aunque el formato de este artículo permite su lectura independiente de su primera parte, el lector puede requerir alguna referencia a términos o ejemplos utilizados en el.¹

Una buena parte de las decisiones que el trabajo clínico diario requiere tienen que ver con la implementación de alguna forma de intervención con la que esperamos modificar el curso natural de una enfermedad o proceso. Sin embargo, con frecuencia nos encontramos con escasa claridad o incertidumbre respecto a la mejor decisión con respecto a alguna intervención, o sus relativas ventajas respecto a su costo o efectos secundarios. ¿A dónde acudir por las respuestas? Y si las encontramos, ¿cómo juzgar su confiabilidad? Y si las encontramos confiables, ¿qué tanto de esto podríamos aplicar? Este artículo presenta una

* Profesor asociado, Facultad de Medicina, Universidad Autónoma de Bucaramanga, Bucaramanga, Colombia. Research Fellow, Population Health Institute; Registered Student, Clinical Health Sciences (Health Research Methodology) Graduate Program, Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canadá.

Correspondencia: Dr. Villar, Hamilton General Hospital, McMaster Clinic – Population Health, 238 Barton St. East, Hamilton ON, L8L 2X2, Canadá. e-mail: juan@ccc.mcmaster.ca

Tabla 1. Clasificación de los niveles de evidencia y grados de recomendación de acuerdo con el tipo de estudio que genera la información acerca del efecto de intervenciones en salud (material cedido por el Dr. DL Sackett, basado en los niveles de recomendación y grados de evidencia del centro de medicina basada en la evidencia de Oxford).

Grado de recomendación	Nivel de evidencia	Tipo de estudio
A	1a 1b	Revisión sistemática de ECA (homogéneos entre si) ECA individual (con intervalo de confianza estrecho)
B	2a 2b 3a 3b	Revisión sistemática de estudios de cohorte (homogéneos entre si) Estudio individual de cohortes / ECA individual de baja calidad Revisión sistemática de casos y controles (homogéneos entre si) Estudio individual de casos y controles
C	4	Series de casos, estudios de cohortes / casos y controles de baja calidad
D	5	Opiniones de expertos basados en revisión no sistemática de resultados o esquemas fisiopatológicos

discusión alrededor de estas preguntas que tienen que ver con la búsqueda, interpretación y aplicación los ECA en la práctica, de modo que esta se adhiera a los principios de la medicina basada en la evidencia.

¿Por qué buscar ensayos clínicos?

Debido a los atributos que les da su estructura, los ECA se convierten en la herramienta típica para la evaluación del efecto de una intervención de salud. Comparados con otros tipos de estudios, la virtud fundamental de los ECA se encuentra en su posibilidad de suministrar respuestas en donde los sesgos se han disminuido al máximo posible. Un estudio *observacional* no permitiría asegurar que los pacientes tratados tuviesen el mismo tipo o estado de enfermedad, la misma actitud o preferencia frente al tratamiento, las mismas cointervenciones (aquellas adicionales a la intervención experimental), etc. Además, tampoco aseguraría que los clínicos indiquen el tratamiento al mismo tipo de pacientes que a los que no lo reciben. De otro lado, un estudio experimental en donde la asignación del tratamiento no se da al azar podría reducir, pero no suprimir la distribución inequitativa de factores de confusión (conocidos o desconocidos al diseñar el estudio) en los grupos a ser tratados y evaluados. Tampoco podría controlar la inmensa cantidad de situaciones que podrían ocurrir a los pacientes tratados durante el tiempo en el cual se espera que se desarrollen (o no se desarrollen) las respuestas esperadas. Cualquiera de estas situaciones produciría una distorsión inevitable de los resultados en una u otra dirección. Los ECA, mediante el uso de dos condiciones fundamentales, la asignación al azar de la intervención activa u otra que le sirve de control y la observación de acciones que impidan conocer el grupo al que cada participante fue asignado, se convierten en el diseño disponible con mayor capacidad para minimizar esos sesgos. La necesidad de la replicación para aceptar en la ciencia un hecho más como válido ha hecho que el resultado agregado de varios ECA que abordan la misma pregunta y arrojan resultados similares

constituyan el nivel superior en la escala de niveles de evidencia sobre intervenciones (Tabla 1).

El desarrollo de un ECA se da en un escenario clínico en el que una población determinada se somete a la prueba de una nueva intervención asignada al azar sobre la cual se entrega un balance del efecto hipotético, al comparar su efecto frente al cuidado convencional o una segunda intervención que le sirve de control. De ahí es que sus resultados se dan términos de diferencias, en donde el sustraendo es el grupo control. En la búsqueda de respuestas en medio de fuertes limitaciones de tiempo, el clínico de cuando en cuando encuentra que un ECA es desarrollado en un escenario y tipo de población dado, y estudia el efecto de la intervención de interés, reproduciendo con exactitud la situación sobre la cual tenía incertidumbre o carecía de información. Esto es, con una frecuencia cada vez mayor, nos encontraremos con ECA que abordan de manera específica nuestras preguntas y pueden representar una fuente válida y valiosa de ayuda. El clínico puede encontrar en los ECA respuesta a preguntas a un variado espectro de intervenciones o escenarios. Aunque con mucho más frecuencia los ECA se utilizan para medir la eficacia de medicamentos, estos no se limitan solo a la prueba de este tipo de intervenciones. De hecho, estos pueden probar intervenciones tan variadas como procedimientos quirúrgicos, aparatos mecánicos, dietas, intervenciones educativas, etc. Mientras sea posible mantener la adherencia a su estructura fundamental, no hay limitaciones. Baste señalar como ejemplo gráfico la realización de un ECA sobre el efecto de las oraciones en la sala de espera sobre los desenlaces de pacientes en la sala de cuidados intensivos.^{2, 3} Aunque ante la prueba de diferente tipo de intervenciones un ECA puede tomar diferentes formas en su conducción, los puntos que destacábamos de su estructura metodológica básica constituyen el *sine qua non* de lo que se acepta como requerido para generar respuestas válidas. Las consideraciones éticas o logísticas, más que la falta de adherencia de la comunidad científica a estos principios

básicos, son las limitantes a la expansión aun mayor de su uso para probar intervenciones. Es claro que desde sus inicios al final de los años 40, los ECA se han desarrollado, expandido y diversificado en forma notoria, si bien son aun una especie imperfecta en constante evolución.⁴ Pero a pesar de sus actuales limitaciones, se acepta que estos constituyen el elemento integrador más avanzado entre la investigación clínica y la implementación de sus resultados en los servicios de salud.⁵ Las respuestas que los ECA proporcionan, en el contexto de las características de los servicios de salud, los valores de pacientes y terapeutas involucrados y -por encima de todo esto- el raciocinio y la experiencia del clínico son la base para la práctica de la medicina basada en la evidencia.⁶

Búsqueda y selección de un ensayo clínico

Con la idea de participar de una manera más informada en el manejo de su paciente, un internista general se prepara para asistir a una junta médica en la que se hablará sobre el plan quirúrgico que se propone para uno de sus pacientes, quien además representa un caso común en su práctica: Un paciente masculino de 67 años con diabetes tipo II e hipertensión arterial en tratamiento, con repetidos ataques isquémicos transitorios desde hace 3 años, a quien se le ha recomendado una cirugía de catarata, con la intención de mejorar su calidad de vida. El plan anestésico recomendado es utilizar anestesia local mediante un bloqueo peribulbar para su cirugía. Usando

como base los resultados de un ensayo clínico reciente,⁷ el plan propuesto es usar una solución en la que se adiciona una mezcla de opioides a los anestésicos locales. A pesar de su escasa familiaridad con la técnica y sus implicaciones directas, nuestro internista desea hacerse a su propia idea sobre esta decisión.

A sabiendas de que no dispone de más de una hora para ello, apela a su conexión a Internet para buscar la evidencia que existe sobre la efectividad y seguridad de esta intervención. Después de consultar a un colega con un poco más de familiaridad con las búsquedas de información, decide visitar tres sitios en la red: La biblioteca Cochrane⁸ y las revistas "secundarias" (donde se seleccionan, condensan y presentan estudios que se consideran especialmente valiosos) que él conoce: ACP Journal y la colección Evidence-based Journals que maneja el grupo editorial del British Medical Journal^{9,10} y Pubmed, el servicio gratuito de búsqueda de la Biblioteca Nacional de Medicina de los Estados Unidos.¹¹

La entrada a la biblioteca Cochrane pone al consultante en alerta de la gran cantidad de recursos que se han condensado allí, los cuales decide imprimir para referencia posterior (Tabla 2). Partiendo de los niveles de evidencia centra su atención sobre la base de revisiones sistemáticas Cochrane y el registro de ensayos clínicos controlados. Usando inicialmente las palabras "*peribulbar anesthesia*" (lo que la biblioteca interpreta como "*peribulbar*" y "*anesthesia*") aparecen 0 revisiones sistemáticas y 116 ECA registrados. Al insistir en una revisión sistemática,

Tabla 2. Recursos disponibles en la biblioteca Cochrane y número de referencias disponibles (a noviembre de 2001) en cada uno de sus componentes. Adaptado de la página inicial en el sitio de Internet de la biblioteca Cochrane (<http://www.cochranelibrary.com/clib/>)

Tipo de base de datos	Tipo de material	Número de referencias
Revisiones sistemáticas Cochrane	Revisiones completas Protocolos	1.235 956
Resúmenes de revisiones de efectividad	Resúmenes de evaluación de la calidad de revisiones sistemáticas	2.208
	Otras revisiones: Solo detalles bibliográficos	800
Registro de ensayos clínicos controlados	Referencias	327.698
Revisiones metodológicas	Revisiones completas	2
	Protocolos	7
Registro metodológico	Referencias	3.571
Valoraciones de tecnología en salud	Resúmenes de evaluación de tecnologías	2.306
Evaluaciones económicas	Resúmenes de evaluación económica de intervenciones	8.086

acepta ser menos específico en la búsqueda y utiliza luego "local anesthesia" y simplemente "anesthesia", lo que genera 3 y 48 revisiones, respectivamente. Una hojeada a los títulos de las revisiones le permite identificar dos de ellas como potencialmente útiles. La primera estudia la comparación entre la anestesia local y la general para endarterectomía carotídea,¹² lo que a pesar de referirse a una intervención diferente podría ser de interés por incluir una población como la que le interesa. La otra se refiere a la efectividad para el dolor postoperatorio de una infiltración anestésica en amigdalectomías,¹³ que se refiere a una población e intervención diferentes, pero que podría presentar información útil por los desenlaces evaluados. Una mirada a los resúmenes le indica que, de un lado, la anestesia local versus general no produjo diferencias en la mortalidad perioperatoria a 30 días, después de sintetizar información de 15 estudios no aleatorizados que incluían 5,970 pacientes (Odds ratio = 0.61, intervalos de confianza del 95% 0.32, 1.16). En la otra revisión sistemática, la síntesis de 5 ensayos clínicos evaluados como de baja calidad metodológica no permitió documentar diferencias en el dolor postoperatorio de los niños sometidos a amigdalectomía. Al volver al registro Cochrane de ECA, una hojeada a los títulos le indica que en apariencia no hay ninguno que evalúe al uso de opioides dentro de la mezcla usada en la anestesia peribulbar. Al abrir algunos de los enlaces con ellos encuentra que no es posible refinar un poco más la búsqueda en la biblioteca Cochrane y recuerda que podría hacer esto en Pubmed, el servicio gratuito de búsqueda de información contenida en la Biblioteca Nacional de Salud de los EEUU, en Internet.

En una base de datos que incluye revistas secundarias (Ovid, McMaster University) en donde se condensa la serie de revistas de medicina basada en la evidencia y el ACP journal, al buscar bajo los términos "*peribulbar anesthesia*" y luego "*local anesthesia*" no se obtienen referencias. Bajo "*anesthesia*" se encontraron en cambio 19 comentarios a publicaciones originales, pero al revisar los títulos ninguna de ellas fue relevante a su pregunta.

A pesar de no encontrar revisiones sistemáticas, nuestro colega para ese momento conocía de la publicación de un centenar de ECA en el tema de la anestesia peribulbar y continuaba a la expectativa de hacer aportes a la discusión del caso en la junta médica. Al explorar Pubmed, inicia la búsqueda usando "peribulbar anesthesia" y obtiene 350 resultados. Luego descubre la opción de límites y refina la búsqueda limitándola por tipo de estudio, seleccionando la opción "randomized controlled trials" y obtiene 99 referencias. Al descubrir la opción "history" observa la posibilidad de combinar búsquedas. De este modo, añade la "peribulbar blockade" al anterior término, limitando su búsqueda a ECA. La combinación de estos términos produce 106 referencias. Dada la necesidad de refinar aun más su búsqueda, decide tratar de asegurarse de revisar la información dada provista por ECA. Para asegurarse de esto, solicita las referencias que contengan las palabras "randomized" or "randomised" (ante la posibilidad de

publicaciones que usen el inglés británico) or "blind", de un lado and "fentanyl" or "morphine" or "lidocaine" or "bupivacaine", teniendo como base los medicamentos a los que se refiere el estudio que se mencionó en esa junta médica. La búsqueda genera ahora nueve referencias (ver el apéndice al final del texto), un número que este colega considera puede revisar con algo más de detalle.

Dada las limitaciones de tiempo y acceso a los materiales completos, decide seguir las recomendaciones que había aprendido para preseleccionar referencias relevantes a su pregunta "en pacientes de alto riesgo para eventos posquirúrgicos sometidos a cirugía de catarata, ¿cual es la eficacia y seguridad del bloqueo peribulbar usando una mezcla de opiáceos y anestésicos locales?" Después de imprimir los resúmenes de las referencias obtenidas, decide elaborar una tabla que le ayude a ponderar la evidencia de importancia para su pregunta (Tabla 3).

Un examen de este resumen de puntos claves de los ECA encontrados en Medline le demuestra que sus criterios de selección de ECA fueron muy específicos: ocho de los nueve estudios fueron en realidad ECA. De estos ocho estudios, siete fueron publicados en inglés y otro en italiano, lo que podría ser atribuible a la subrepresentación de fuentes de información publicada en otros lenguajes en Medline (14; 15). Sin embargo, el hecho de encontrar estudios cuyo autor principal estaba afiliado a instituciones de Italia (n = 3), el Reino Unido (n = 2), Alemania (n = 1), India (n = 1) y Estados Unidos (n = 1) le permite inferir que el interés por la investigación en este tópico es universal. Siete de los nueve estudios de la lista son publicados a partir de 1999, lo que sugiere que el interés es actual, aunque plantearía la contraparte de la falta de conocimiento de los efectos post implementación de este tratamiento, los llamados ensayos clínicos de fase 4,¹⁶ tal como nuestro colega lo recuerda con el caso de los efectos secundarios de algunas estatinas, en donde la evidencia sobre sus efectos secundarios se hizo evidente después de varios años de uso masivo.^{17,18}

Llama la atención de nuestro colega el hecho de que el título de sólo tres de los estudios resumidos en la tabla de referencias extractadas informa el carácter aleatorio y el tipo de enmascaramiento usado para ocultar la intervención asignada. De ahí que los expertos reunidos en el consenso CONSORT hayan pedido que mejore la calidad de los informes de ensayos clínicos y recomienden unos criterios para evaluar su calidad.¹⁹ Idealmente para un usuario regular, tal como nuestro colega, un ECA debería poder ser identificado de forma rápida a partir del título. Esto implica que se pueda reconocer al estudio como un ECA y que se describan la intervención y población estudiada, así como los desenlaces evaluados. Usando como ejemplo el título del estudio de Rangel y cols que se ha sugerido seguir para tomar decisiones en nuestro paciente hipotético ("Bloqueo peribulbar adicionando opioides en analgesia posoperatoria para cirugía oftálmica. Un ensayo clínico controlado"), éste informa de manera apropiada sobre varios de estos elementos. Quizá un título

Tabla 3. Principales características encontradas en el resumen de 9 ensayos clínicos relacionados con la eficacia del uso de bloqueos peribulbares en cirugía oftálmica identificados en Medline de acuerdo con la estrategia descrita en el texto.

Ref	Autor País – año Idioma	Titulo informa		Población elegible, Unidad aleatorizada (asignados a tratamiento)	Intervención grupo activo (A) vs control (C)	Desenlace(s)	Resultado general
		Aleatorio	Ciego				
1	Agrawal India- 1993 Inglés	Si	Si	Cirugía de catarata, pacientes (122)	A- Bloqueo peribulbar (NE) C- Bloqueo retrobulbar (NE)	Anestesia globo ocular (1) Tiempo de latencia (ojo y párpado) Dilatación pupilar Hipotonía ♦	Neutro
2	Anders Alemania-1999 Inglés	Si	No	Cirugía de catarata, Ojos (200)	A- Anestesia intracameral (lidocaína + anestesia tópica con oxybuprocaina) C- Anestesia retrobulbar (prilocaina)	Complicaciones postoperatorias Tiempo quirúrgico, Agudeza visual posquirúrgica Latencia ondas B del electroretinograma	Neutro
3	Azuara-Blanco EEUU-1997 Inglés	Si	No	Trabeculectomía, Pacientes(60)	A- Anestesia subconjuntival (4 ml de una mezcla 1:1 de mepivacaína 2% + bupivacaína 0.75% + hialuronidasa) C- Anestesia peribulbar (2 ml de la mezcla)	Dolor intraoperatorio Aguinesia ocular Complicaciones intraoperatorias	Neutro
4	Bedi Reino Unido-1999 Inglés	No	Si	Cirugía de catarata, Pacientes(60)	A1- lignocaína 2% + adrenalina A2- prilocaína 3% + felipressina 0.03 UI/ml A3- lignocaína 2% + bupivacaína 0.5% Todas las mezclas en proporción 1:1	Dolor a la inyección Tiempo de latencia del bloqueo Condiciones operatorias iniciales y finales ♦	Dolor >A1 Latencia neutro Aguinesia >A2
5	Costa Italia-1999 Italiano	Si	No	Cirugía de catarata, Pacientes(399)	A- Bloqueo peribulbar A1- anestésicos locales + hialuronidasa A2- bupivacaína C- Bloqueo retrobulbar C1- anestésicos locales + hialuronidasa C2- bupivacaína	Período de latencia Movimientos oculares residuales Necesidad de anestesia adicional Calidad de la anestesia ♦ Tono ocular Longitud de la anestesia Complicaciones ♦	Mejor A1
6	Davis Canadá-1989 Inglés	No	Si	Cirugía de catarata ** (90)	Hidrocarbonato de Lidocaína 2% Hidrocloruro de Lidocaína (hialuronidasa agregada en ambos casos)	Aguinesia a los 4' de la inyección	Neutro
7	Gioia Italia-1999 Inglés	No	Si	Cirugía vítreo-retiniana ASA I o II(60)	Bloqueo peribulbar con A- ropivacaína 0.75% 8 ml C- lidocaína 2% + bupivacaína 0.5%	Tiempo de anestesia Calidad de la anestesia posoperatoria + Efectos secundarios Consumo de analgésicos	Neutro
8	Luchetti Italia-2000 Inglés	Si	Si	Cirugía de catarata (2000)	Bloqueo peribulbar con A- ropivacaína 0.75% 9ml C- bupivacaína 0.5% 4 ml + mepivacaína 2% 4ml (hialuronidasa agregada en ambos)	Dolor a la inyección Tiempo de latencia Necesidad de dosis adicionales Tiempo de anestesia Analgésia intraoperatoria Duración/Complicaciones cirugía	Mejor A
9	Perello Reino Unido-2000 Inglés	Si	Si	Cirugía de catarata (90)	Bloqueo peribulbar con A- ropivacaína 0.75% 9ml C1- ropivacaína + bupivacaína C2- lidocaína + bupivacaína	Tiempo de latencia Satisfacción quirúrgica ♦ Satisfacción del paciente ♦ Complicaciones y efectos secundarios	Neutro

** Asignación de intervenciones no aleatoria; (1) Señalado en el resumen como el desenlace primario del estudio; ° Variables no especificadas; + Se presenta en el resumen del estudio que 26/30 pacientes aleatorizados a A y 18/30 pacientes asignados a C informaron ausencia de dolor al primer día posoperatorio (p=0.005). Nuestro cálculo para esas proporciones señala un valor de p= 0.041.

aún más preciso evitaría que usuarios interesados, por ejemplo, en el efecto de esta intervención en poblaciones diferentes a la geriátrica seleccionaran este ECA, que podría no ser de su interés. De la misma manera, el uso del término “cirugía oftálmica” deja todavía una base amplia que el de la cirugía de catarata (el término usado en el resumen e incluido como una de las palabras claves) con la que algunos usuarios con otros intereses podrían seleccionar este ECA.

Desde el punto de vista del tamaño de su población, todos los ECA presentados en la tabla 2, excepto uno, son pequeños ($n < 500$). Aunque este número de participantes puede ser suficiente para demostrar efectos importantes en la clínica, en especial algunos desenlaces relacionados con variables continuas medidas en todos los participantes (v.g. el tiempo de latencia del bloqueo peribulbar), puede ser insuficiente para la detección de eventos infrecuentes (v.g. complicaciones) o aparecidos en el largo plazo (v.g. fibrosis de la grasa periorbitaria). Aunque cada una de estas consideraciones de tamaño son relativas al problema particular que un ECA esté abordando, no existe discusión acerca de que los ECA pequeños tienen una mayor probabilidad de producir estimados del efecto de la intervención falsamente “positivos” (con errores alfa) y menos precisos (con intervalos de confianza más amplios). Es decir, la estimación del efecto (dado que el efecto “real” no será nunca determinado, sino estimado) tiene una probabilidad mayor de desviarse de la realidad. Además, los ECA de pequeño tamaño tendrían mucho mayores problemas para la generalización de sus resultados.²⁰ Aunque la población estudiada en estos ECA está constituida fundamentalmente por pacientes candidatos a cirugía de catarata, el efecto de la intervención propuesta para este paciente -la adición de opiáceos a la mezcla de anestésicos locales- no fue evaluada en ninguno de ellos.

A pesar de que los ECA encontrados no abordan específicamente la pregunta sobre la cual se deseaba obtener información (como ocurre aun con mucha frecuencia), un listado de características de los ECA en el área, como el de la tabla 3 puede ser de utilidad. Por ejemplo, la lista de los desenlaces evaluados en esos ECA denota los asuntos que se consideran un problema que debe resolverse para los investigadores en este tópico. En esta lista el periodo de latencia (siguiendo la definición usada por Rangel y cols) aparece como desenlace en seis de los nueve estudios y el tiempo de analgesia posquirúrgica aparece en cuatro de estos. Estos desenlaces parecen ser los que más preocupan a los autores; por tanto, deberían ser elementos centrales en la toma de decisiones sobre la adopción de la nueva intervención que se propone. El estudio de Rangel coincide con esas apreciaciones y mide esos desenlaces. La otra observación posible en esta lista de ECA es que en la comparación de la eficacia de diferentes mezclas o agentes de anestésicos locales o sus vehículos y coadyuvantes, la mayoría de los estudios encuentra un efecto neutro (v.g. no diferencia) en las comparaciones estudiadas. Este hecho puede indicar la

búsqueda de alternativas para encontrar la mejor combinación y mejorar el cuidado de salud en esta área. En ese contexto, la pregunta formulada por Rangel es muy plausible.

Pasada esta única hora disponible para dar una rápida mirada a la literatura disponible, nuestro colega internista ha encontrado que al menos en las fuentes más consultadas de información sobre la eficacia de intervenciones, la pregunta específica no ha sido abordada mediante ECA, excepto en el estudio que sirve de argumento a esta decisión. Conoce que la técnica anestésica no ha sido probada por largo tiempo, ni en grandes poblaciones y que, por lo tanto, no hay información acerca de la incidencia de complicaciones infrecuentes pero de alto costo o surgidas en el largo plazo. También se ha enterado de cuáles son los desenlaces que más preocupan a quienes hacen la técnica. Como resultado de este proceso de búsqueda, nuestro colega ha tomado la decisión de no leer ninguno de los artículos cuyos resúmenes revisó y más bien concentrarse en el análisis de la validez y aplicabilidad del estudio de Rangel y cols. Ha encontrado que el criterio fundamental para seleccionar un ECA para posterior lectura y evaluación, el de la importancia de la pregunta que aborda en el contexto del cuidado de salud en esa área, se cumple a cabalidad en este estudio.

Evaluación de un ensayo clínico

Una vez seleccionado un ECA por su ayuda potencial para resolver nuestras preguntas de interés práctico, la curiosidad nos lleva inmediatamente a tomar nota de cuales fueron la dirección (positivo, neutro o negativo), la magnitud (una mayor o menor diferencia entre el grupo de tratamiento activo y el control) y la precisión (el ancho de los intervalos de confianza de esas diferencias) de los resultados de los ECA. Sin embargo, siendo el responsable de su paciente, el clínico prefiere antes de tomar decisiones con base en esos resultados, abordar un paso previo y necesario: el análisis de la validez del estudio, cualesquiera sean sus resultados.

Existe consenso en que la calidad metodológica de los ECA consiste precisamente en la observación de los principios que le son propios discutidos en otras partes de este artículo. Aunque el peso que diferentes metodólogos dan a cada componente puede variar, todas las recomendaciones apuntan al examen de la aleatorización, el mantenimiento del enmascaramiento de las intervenciones para los diferentes actores que toman parte en el estudio y la observación de un seguimiento en el máximo número posible de participantes como los elementos claves en la evaluación de la calidad de un ECA. Así se observa en la lista de criterios del grupo de trabajo de medicina basada en la evidencia para la evaluación de estudios sobre prevención o tratamiento que se presenta en la tabla 4.²¹

Para la evaluación del ECA publicado por Rangel y cols en la reunión, se utilizaron estos criterios y algunas referencias adicionales: La escala de Jadad (Tabla 5),²²

Tabla 4. Lista de criterios establecidos por el grupo de trabajo de medicina basada en la evidencia (McMaster University) para la evaluación de la validez de un estudio sobre terapia o prevención. Adaptada de la referencia 21.

Criterios primarios	¿Se asignaron aleatoriamente los tratamientos a los participantes?
	¿Se incluyeron en el análisis todos los pacientes hasta la finalización del período de seguimiento?
	¿Se realizó un seguimiento completo?
	¿Se analizaron los pacientes como parte de los grupos a los que fueron inicialmente asignados?
Criterios Secundarios	¿Se mantuvo el enmascaramiento del tratamiento asignado a los diferentes grupos de personas involucrados en el estudio (participantes, operadores o clínicos, quienes adjudican los eventos de interés, analistas de datos y personal del estudio)?
	¿Fueron similares los grupos al inicio del estudio?
	Además de la intervención experimental, ¿se aplicaron las mismas intervenciones a los grupos que están siendo comparados?

una herramienta validada, de aplicación simple y a la que se recurre con gran frecuencia, y los métodos utilizados para la valorar la calidad y seleccionar los ECA en las dos revisiones Cochrane que nuestro colega había antes extractado. Ya en la reunión se presentó a discusión el caso clínico y el artículo de Rangel y cols, con la intención de discutir su potencial aplicación, como se había sugerido. Se reconoció a este ECA como novedoso en el área, de impacto potencial sobre la práctica y en particular, se destacó su realización en un medio sin mayores precedentes en este tipo de estudios. El asunto de la discusión de la validez de este ECA en particular era crítico, pues se trataba del único disponible para ilustrar la decisión clínica. Había tanto prevención como interés sobre el tema de la calidad cuando se expuso que en las dos revisiones Cochrane después de la evaluación de la calidad, la mayoría de estudios fueron excluidos. En la primera de ellas,¹² de 20 estudios identificados, 17 no asignaron aleatoriamente la intervención en estudio. En la segunda,¹³ de 30 ECA identificados, sólo seis cumplieron con todos los criterios de inclusión en esa revisión. Estos fueron: 1) Uso de un método de aleatorización adecuada; 2) Inclusión de un número adecuado de participantes por grupo de estudio (>10); 3) Minimización de los sesgos posteriores a

la asignación aleatoria (pérdidas del seguimiento y análisis por intención de tratamiento); 4) Preservación del enmascaramiento de quienes evaluaban los desenlaces en los participantes; y, 5) Calidad de la medición de los desenlaces (v.g., escala análoga de dolor validada).

El ECA de Rangel y cols fue presentado como aleatorizado, controlado y triple ciego en el informe, lo que sugería a los participantes en la reunión una buena calidad. El estudio evaluó tres grupos asignados a tratamiento, dos activos (agregando morfina o fentanyl a una mezcla de anestésicos locales) y un control que correspondía a la práctica convencional (la solución de anestésicos locales). La asignación fue hecha por medio de una tabla de números aleatorios, un método perfectamente aceptable. Sin embargo, no se especificó cómo se usó la tabla para generar la secuencia y establecer la asignación (por ejemplo en grupos de números, o números precodificados). Desde el punto de vista del tamaño de los grupos (@ 30 por grupo), la posibilidad de desbalance en la prevalencia de variables de confusión entre los grupos era intermedia. Si bien los valores de p para la comparación de las características entre los grupos fuese >0.05, la declaración de ausencia de diferencias en estas comparaciones no tenía un alto

Tabla 5. Escala validada para la valoración de la calidad de los ensayos clínicos controlados propuesta por Jadad, en la que se asignan o restan puntos a cada criterio. Un ensayo clínico de buena calidad tendría más de 3 puntos. Adaptado de la referencia 22.

	Asignación aleatoria de las intervenciones		Enmascaramiento de la intervención asignada durante el estudio		Conteo de las pérdidas durante el seguimiento de los participantes	
¿Descrita en el informe?	Si	+1	Si	+1	Si	+1
	No	0	No	0	No	0
¿Apropiada?	Si	+1	Si	+1		
	No	-1	No	-1		

poder. De otro lado, el uso de tres grupos disminuía aún más este poder. En conclusión sobre este punto, los participantes en la junta médica coincidieron en que si bien se usó un método adecuado para la asignación, el tamaño y el diseño del estudio no podría generar una completa confianza sobre la distribución balanceada de variables de confusión en los diferentes grupos de estudio, lo que habría de tenerse en cuenta para la valoración de los resultados.

Algunos aspectos derivados de la intervención experimental fueron también tratados. La descripción en el informe demuestra que se tuvo la precaución de aparear las mezclas experimentales por el volumen a ser infundido, como potencial sesgo posaleatorización. No se describieron las cointervenciones en el acto operatorio o el posoperatorio inmediato que pudieran haber modificado los resultados. Como ejemplo de sesgo posaleatorización se mencionó la interacción que fue documentada los estudios para evaluar intervenciones para evaluar la prevención de la muerte súbita por arritmias en pacientes posinfarto en la que se comparaba la eficacia de la amiodarona (un antiarrítmico) con los cardiodesfibriladores implantables (un dispositivo).²³ En estos ECA una frecuencia de uso de betabloqueadores en los grupos de estudio posterior a la aleatorización produjo efecto sobre los resultados al final de los ECA. Dado que el informe de Rangel no presentó detalles sobre cointervenciones durante el seguimiento de los pacientes en esas 48 horas de evaluación del dolor postoperatorio, se generaron algunas dudas sobre una eventual modificación de los resultados.

Se discutió entonces el tema del mantenimiento del enmascaramiento dentro de este ECA. Por sus efectos potenciales sobre los resultados, éste fue el punto de mayor discusión en la reunión. Se recordó que una revisión sistemática de ECA que evaluaban intervenciones para el dolor, el mantenimiento del enmascaramiento resultó la característica metodológica más importante para la obtención de resultados válidos.²⁴ Por el carácter de la intervención en cuestión, un procedimiento anestésico en donde el factor experimental son los componentes de la mezcla, el enmascaramiento del operador (el anesthesiólogo) para la intervención era importante. El informe no describe tal enmascaramiento. Allí se describe que la enfermera de salas de cirugía preparó las mezclas, pero no es claro de qué forma la enfermera conoció la asignación de cada paciente y de qué forma el operador fue enmascarado para el tipo de intervención utilizada. Hay consenso entre los expertos en la necesidad del placebo usado en los ECA tengan las mismas características físicas de la intervención experimental.²⁵ Sin embargo, esto no siempre es logísticamente posible, como en el ejemplo de los cardiodesfibriladores implantables en donde es inocultable la intervención experimental para varias de las personas involucradas en el ECA. A este tipo de ECA se les ha llamado de código abierto (*open label*). Para el caso del ECA en discusión, se reconoció que en este caso podría haber sido posible un esfuerzo para ocultar la intervención asignada (el tipo de mezcla, más

que el procedimiento anestésico) al anesthesiólogo e idealmente a otras personas que también intervienen en diferentes formas en los desenlaces (v.g. los cirujanos o el personal de la sala de cirugía). La falta de esta característica podría haber influido en los resultados de este estudio.

Fueron evaluadas entonces las diferentes instancias susceptibles de enmascaramiento de la asignación de las intervenciones en un ECA. Los pacientes, los operadores, quienes adjudican los eventos, quienes analizan la información y quienes escriben los informes. Este enmascaramiento resulta aún más importante cuando varias de estas funciones son asumidas por las mismas personas (por ejemplo, que los operadores sean a su vez los investigadores). El informe de este ECA lo presenta como estudio triple ciego debido a que ni los pacientes, ni quienes adjudicaron los desenlaces, ni quienes analizaron los datos conocían la intervención asignada. En la reunión alguien comentó que eso no era lo que se entendía por triple ciego. No hubo acuerdo después de unos minutos de discusión.

El uso de estos términos está muy arraigado en la terminología de los ECA, particularmente el uso del “doble ciego”, tal como se puede ver por la evolución del número de referencias registradas en Medline que usan estos términos (Tabla 6). Sin embargo, así como a nuestros colegas de la reunión, no hay tampoco acuerdo entre los clínicos en otros centros y países, ni siquiera entre los investigadores que conducen ECA, ni entre los expertos en metodología. Un estudio evaluó el acuerdo alrededor de lo que se entendía por los términos “single-blind”, “double-blind” y “triple-blind” en los ECA en tres diferentes escenarios: a) Entre 91 internistas generales en tres diferentes centros de Canadá; b) En una muestra de 200 ECA, que correspondía a los 40 más recientes ECA encontrados sistemáticamente en 5 revistas (*Annals of Internal Medicine*, *BMJ*, *JAMA*, *The Lancet* y *The New England Journal of Medicine*) y c) En 25 textos publicados después de 1999, identificados sistemáticamente. Los resultados demostraron que con el tiempo la consistencia fue menor en las tres fuentes consultadas para el acuerdo en la definición de ciego simple, doble y triple para cada caso. La interpretación más común entre médicos y textos para ciego simple fue 75% y 74%; doble 38% y 43% y triple 18% y 14%. Para el caso del triple ciego, los médicos identificaron 15 interpretaciones diferentes y los textos presentaron 7 definiciones. Ningún estudio de la muestra de 200 ECA fue identificado como triple ciego. La discrepancia en la definición de ciego simple, doble o triple tendría importancia solo en la medida en que el informe de un ECA no se especifique cuál de las personas implicadas se mantuvo enmascarada para la asignación de intervenciones. Entre los 200 ECA publicados que fueron evaluados en el estudio mencionado arriba, 83 fueron descritos como doble ciego, pero en 41 de ellos no se especificó cuáles eran los grupos enmascarados. En el informe de Rangel y cols se describieron de manera clara quienes fueron los grupos enmascarados.

Tabla 6. Número de referencias obtenidas en Medline, en tres diferentes períodos de tiempo, restringiendo la búsqueda a ensayos clínicos controlados y a los siguientes términos, si estaban contenidos en el título o resumen: (1) "Open label"; (2) "Double-blind"; (3) "Triple-blind"; (4) "Quadruple-blind". Datos tomados del sitio del servicio PubMed (<http://www.ncbi.nlm.nih.gov/PubMed/>)

	1967-1980	1981-1990	1991-2001	Total
Código abierto (1)	4	121	1.492	1.617
Doble ciego (2)	2.497	10.907	24.626	38.030
Triple ciego (3)	9	10	27	46
Cuádruple ciego (4)	0	0	1	1
Total	2.510	11.038	26.146	39.694

Pero aun así cuáles son las implicaciones para la validez de este ECA al enmascarar estos tres grupos? Como se discutió arriba, en el informe de Rangel y cols. no se describió que los operadores estuviesen enmascarados. Varios de los desenlaces intraoperatorios y posoperatorios inmediatos fueron evaluados probablemente por los operadores. Entre ellos está el tiempo de latencia, uno de los desenlaces en el que se encontraron diferencias significativas entre los grupos activos y el control. En este caso específico el enmascaramiento del operador tiene una importancia relativa mucho mayor que el de los participantes por razones obvias. Para otro desenlace importante, el tiempo de analgesia posoperatoria, quienes adjudicaron el dolor posoperatorio y en respuesta a esa evaluación administraron analgésicos en la sala de recuperación fueron descritos como enmascarados, lo que en teoría produciría una evaluación más libre de sesgos. Sin embargo, este segundo tipo de desenlace podría ser visto como una consecuencia mediata del procedimiento anestésico cuyo operador no había sido enmascarado. En ese contexto en donde un desenlace es en realidad un efecto mediato de una intervención hecha con código abierto, el enmascaramiento del segundo evaluador puede considerarse, al menos parcialmente, influido por ese código abierto, por lo tanto menos influyente en los resultados. Para este tipo de ECA el enmascaramiento tendría jerarquías, siendo la más importante la del operador. Una vez adjudicados los desenlaces por personal que no está completamente enmascarado y generado un "sesgo por conocimiento" (*ascertainment bias*), el efecto del enmascaramiento de otro tipo de personas para reducir este sesgo (v.g. el analista de datos) sería mínimo, aunque sea desde luego deseable el enmascaramiento del mayor grupo posible de personas alrededor del ECA. Este argumento aplicaría incluso al enmascaramiento de quien escribiera el informe, si fuese una persona diferente al operador. Así, después de alguna discusión sobre el significado del ciego triple en este ECA, los asistentes coincidieron en que los resultados de los desenlaces, en especial los desenlaces operatorios inmediatos, debían interpretarse a la luz de esas precauciones.

El seguimiento de este ECA se realizó por 48 horas, con el objetivo de evaluar el efecto analgésico posoperatorio de la muestra. No se diseñó este ECA para observar el efecto

de las diferentes mezclas en el mediano o largo plazo. Esto generó preocupación en nuestro colega internista, por el tipo de pacientes que por lo general se someten a cirugía de cataratas y se les indica anestesia local para los procedimientos. En esa población la comorbilidad es la norma y esta preocupación tiene sentido. De hecho, los participantes del ECA evaluado pertenecían a la clase ASA II en un 69.7%, 72.4% y 61.1% en los tres grupos de estudio. Se argumentó entonces que desde la óptica de un anestesiólogo, los desenlaces importantes son en principio los derivados del acto anestésico de una forma más directa. En el estudio de Rangel y cols. se informó una mayor frecuencia de complicaciones del procedimiento anestésico. No se especificó o se comentó en la discusión su significado potencial, lo que dejó esta pregunta abierta. Se revisó entonces la distribución de las frecuencias del tiempo de analgesia, el segundo de los desenlaces en donde se observó una diferencia significativa. Se recordó la discusión anterior sobre el enmascaramiento y, haciendo esa salvedad, hubo consenso acerca de que este resultado considerado positivo debía sopesarse con las complicaciones del procedimiento para una eventual recomendación del uso de esta mezcla en este paciente.

Un seguimiento suficiente para el desarrollo de los desenlaces esperados es otro de los aspectos críticos en la evaluación de los ECA. El ECA publicado por Rangel y cols. realizó un seguimiento de 48 horas y evaluó el tiempo de analgesia a lo largo de este período. En el marco de este lapso de seguimiento, este ECA pudo presentar información en 98 de 104 paciente aleatorizados (pérdidas de 5.7%). Se reconoció como muy importante el esfuerzo de los autores por hacer un seguimiento y registrar la gran mayoría de información requerida.

En las tablas del informe de Rangel y cols se presentan 29 comparaciones en donde se calcula un valor de p. Establecido un valor de significancia de 0.05 como demostrativo de una diferencia real entre dos o más grupos, significa esto que por el juego del azar 1 de cada 20 (5%) de las comparaciones podrían encontrar una diferencia inexistente (error alfa). Se quiso entonces evaluar la fortaleza (*robustness*) de las diferencias "positivas" encontradas en este ECA. Esta evaluación se refería al análisis por intención de tratamiento, asignando el mejor

y peor escenario a las pérdidas. Aunque en principio se dijo que unas pérdidas tan pequeñas no debían incidir en los resultados, uno de los colegas anesthesiólogos muy interesado en la analgesia posoperatorio en las primeras 6 horas insistió en discutir esto y presentó las proporciones de la tabla 5 del informe de Rangel y cols. de una manera diferente. Al considerar las pérdidas (un paciente del grupo A, dos del B y uno del C) y reasignar las celdas asumiendo el peor escenario estas diferencias se diluían al categorizar las respuestas dos categorías antes y después de las seis horas (Tabla 7). Se concluyó que aunque las pérdidas eran en apariencia pequeñas, estas tenían el potencial de afectar los resultados.

Luego de esta evaluación de la validez y la sensibilidad de los resultados, el análisis se concentró en su magnitud y precisión. Se destacó que el informe del ECA señalaba con claridad la información de los desenlaces y la prueba de hipótesis de no diferencias entre los grupos de tratamiento. Alguno de los asistentes preguntó si había diferencias entre las dos intervenciones experimentales entre sí, pues no se consignaron en el informe a pesar de ser enunciadas en la sección de métodos. Concentrándose en los desenlaces en los que se encontraron diferencias "positivas", se calcularon las diferencias entre pares de grupos, en donde se confirmó que la diferencia entre cualquiera de las dos soluciones experimentales (B o C) y la solución control (A) eran significativas ($F = 4.19$, $p = 0.044$ para A vs B y $F = 7.67$, $p < 0.001$ para A vs C en EpiInfo 6.04), pero no para la comparación entre los dos tratamientos experimentales ($F = 2.43$, $p = 0.126$ en EpiInfo 6.04). Una situación similar ocurrió con la distribución de los tiempos de analgesia postoperatoria ($X^2 = 2.36$, $p = 0.500$ para la distribución de B vs C usando las categorías de la tabla original). No hubo entonces claridad en la reunión sobre la existencia de diferencias entre la adición de morfina o fentanyl en la mezclas y se recomendó a los investigadores el estudio de esta pregunta en el futuro utilizando un poder mayor para detectar estas diferencias.

En el ECA de Rangel y cols, la muestra fue calculada con base en la consideración de que una diferencia clínicamente significativa para el tiempo de latencia era mayor de un minuto (para dos grupos y estableciendo una desviación estándar de un minuto, como se comentó en la primera parte de este artículo). Así las cosas, aunque la dispersión de los tiempos fue mayor a uno, las diferencias ponderadas (por el tamaño de cada grupo) fueron entre los grupos A y B 1.87 seg (IC 95% 0.06, 3.68; $p = 0.04$ en Meta View. 4.1) y entre los grupos A y C 3.85 seg (IC 95% 2.52, 5.18 en Meta View). Respecto al otro desenlace importante, se discutió la necesidad de no sólo establecer la presencia o ausencia de diferencias, sino la de estimar su magnitud y precisión. Además, estas medidas permiten sintetizar el concepto de eficacia siguiendo los términos de los ECA, la diferencia entre el control y el tratamiento. De acuerdo con el interés que se había manifestado antes, se trató de establecer las diferencias en el tiempo de analgesia posquirúrgica a las seis horas. Para dar un marco de referencia conceptual, se presentó un resumen de las medidas que se usa con más frecuencia para la estimación del efecto de intervenciones (Tabla 8).

Volviendo al ECA de Rangel y cols., la comparación entre las mezclas A y B para las categorías de dolor antes de las 6 horas en forma dicotómica (sí vs no) tal como se presentan en la tabla 7 produciría una reducción del riesgo relativo de 90% (IC 95% desde el 42% al 100%, $p = 0.008$ en Epi Info 6.04). En contraste, la comparación entre A y C arroja una reducción no significativa del 58% (IC 95% desde -20%, en realidad un aumento del riesgo, hasta una reducción del 92% en Epi Info 6.04). Basado en los mismos datos, el número necesario a tratar (NNT) para la intervención B sería 33.3. El NNT es el inverso de la diferencia de riesgos absolutos entre las intervenciones (33.3% para el control y 3.4 para la mezcla B) y es una medida de síntesis del efecto de una intervención fácil de usar y recordar. La interpretación de este número sería que deben tratarse 3.3 pacientes con la mezcla B por cada paciente que mantiene la analgesia posoperatoria por al

Tabla 7. Ejemplo de un análisis de la fortaleza de los resultados de un ECA usando como ejemplo los datos contenidos en el informe del estudio de Rangel y cols (referencia 7). Se muestran arriba (a) los resultados originales (que excluyen a los participantes perdidos del seguimiento) y luego (b) asumiendo el peor escenario, en el que se reasignan esos participantes a las categorías opuestas a la diferencia observada originalmente. Se usaron categorías diferentes a las del informe original, dividiendo el tiempo de analgesia posoperatorio en antes de 6 horas y entre 6 y 48 horas.

	GRUPO A	GRUPO B	GRUPO C	X^2	P
a) Situación original					
Dolor antes de las primeras 6 horas	11	1	5	10.09	0.006
Dolor 6-48 horas	22	28	31		
Pérdidas del seguimiento	1	3	2		
b) Peor escenario					
Dolor antes de las primeras 6 horas	11	4	7	4.16	0.124
Dolor 6-48 horas	23	28	31		

Tabla 8. Resumen de las medidas de eficacia más comúnmente utilizadas para evaluar el efecto de una intervención. Se utiliza aquí el ejemplo del ensayo clínico hipotético presentado en las figuras 1 y 2 de la primera parte de este artículo (referencia 1). * Los cálculos están hechos con el criterio de la intención de tratamiento, usando el número de participantes asignados originalmente a cada tratamiento en el denominador.

	Tratamiento experimental (TE)	Tratamiento control (TC)
Asignados	5.000	5.000
Perdidos del seguimiento *	140	100
Conteo de eventos	400	600
Riesgo absoluto de eventos (RA) # eventos / # asignados	8%	12%
Riesgo relativo de eventos (RR) RA TE / RA TC (IC 95%)	0.67 (0.59, 0.75)	
Reducción del riesgo absoluto (RRA) RA (TE) – RA (TA)	4%	
Reducción del riesgo relativo (RRR) RR * 100 (IC 95%)	33% (25%, 1%)	
Número necesario a tratar (NNT) 1 / RRA	25	

menos 6 horas sin que sea necesaria la administración adicional de analgésicos. El “punto” estimado (el resultado que calculamos) para este efecto debe considerarse bastante alto. Sin embargo, su precisión no es satisfactoria, pues los intervalos de confianza fueron demasiado anchos, en particular para la solución C, en donde estos pasaban el límite de la reducción del riesgo, pudiendo incluso llegar a ser adversas. La solución C entonces no podría ser recomendada para este desenlace. Como conclusión del análisis de los resultados, haciendo caso omiso de la discusión sobre su validez, la junta médica estuvo de acuerdo en que el estudio mostraba que la solución B tenía un efecto favorable sobre ambos desenlaces, aunque sin mucha precisión en la estimación de este efecto. En ese orden de cosas, para el uso de la solución con la adición de morfina se generaría una recomendación grado B.

Un aspecto final de la evaluación de este ECA era su potencial para su aplicabilidad en la práctica general. El hecho de que este ECA evaluara un procedimiento anestésico realizado por un número limitado de operadores afiliados a un solo centro resulta en fortalezas y debilidades de este estudio también. En el centro en donde se realizó, la intervención se había venido practicando con antelación, lo que se supone disminuye la variabilidad interoperador y da una mayor confianza para la aplicación de los resultados allí mismo. Pero a la vez, este hecho podría generar una subestimación de las complicaciones

(de por sí más frecuentes en el grupo experimental en el informe) y quizá una sobreestimación de su efecto benéfico.

De otro lado, este estudio no permitiría la evaluación de la modificación del efecto de la intervención por diferentes tipos de operadores o centros que podrían hacer el procedimiento. Esta falta de replicación en otros escenarios impide la posibilidad de generalizar sus resultados. Se argumentó también que esta situación era el resultado obvio que afrontaba un ECA novedoso como este. Hubo entonces acuerdo en que una recomendación de uso amplio para esta técnica debería esperar. De aplicarse una recomendación usando sus resultados, ésta debería aplicarse exclusivamente en el medio en el cual se desarrolló este ECA.

Aplicación de los resultados de un ensayo clínico

Tomar decisiones con base en ensayos clínicos será el resultado de una ecuación que tiene en cuenta todas las variables mencionadas arriba. La evidencia será un factor que alimente un sistema en el que juegan papel el contexto clínico, junto a los valores de la sociedad, el sistema de salud, los médicos y pacientes. Todos estos elementos se unen a la experiencia y el juicio del clínico, quien tiene a cargo la responsabilidad del resultado final frente a sí mismo, la familia y la sociedad.

En este caso, luego del debate de los puntos mencionados arriba en la evaluación del único estudio disponible para ilustrar la decisión clínica, las conclusiones respecto a los aspectos generales de este ECA: a) Que es el primer estudio conocido que aborda esa pregunta, la cual se considera clínicamente importante; b) Que por lo tanto se carece de información replicada por distintos grupos, así como de información sobre eventos infrecuentes o efectos en el mediano y largo plazo; c) Que se trata de un ECA de pequeño tamaño, de diseño paralelo, en el cual se evaluaron dos intervenciones experimentales frente a un control; d) Que es un ECA de código abierto para los operadores, quienes jugaron papel en la asignación de al menos parte de los desenlaces de su intervención; e) Que sólo algunos de los desenlaces del estudio fueron evaluados en condiciones de enmascaramiento de la intervención; f) Que de acuerdo con los desenlaces que se planeo evaluar en este ECA, se realizó un seguimiento de 48 horas, el cual se completó en aproximadamente el 95% de la población del estudio; g) Que fue un estudio realizado en un solo centro, en el cual se tenía experiencia previa con la técnica.

Respecto a los resultados del ECA se dijo que: a) La adición de morfina a la mezcla de anestésicos locales para anestesia peribulbar en pacientes no seleccionados candidatos a cirugía de catarata produjo un tiempo de latencia más corto y un menor dolor en las primeras 12 horas del posoperatorio, en términos de la escala análoga de dolor. Los resultados para la mezcla con fentanyl son aún imprecisos y no permiten un juicio sobre su eficacia en la analgesia postoperatoria; b) Que como contraparte a la eficacia de la mezcla adicionando morfina, se observó una mayor frecuencia de complicaciones del procedimiento anestésico; y c) Que los números son muy pequeños y el seguimiento muy corto para evaluar las complicaciones postoperatorias menos frecuentes, mediatas o tardías.

Con base en estos elementos se planteó la necesidad de transmitir al paciente y su familiares los pros y contras de la evidencia disponible para tomar una decisión final entre el anestesiólogo tratante y el paciente en una futura cita. Los criterios para la implementación fueron: a) Que la evidencia más sólida generaba una recomendación grado B para la analgesia posoperatoria en las primeras 12 horas; b) Por tanto, este sería el desenlace más importante para generar la decisión de utilizar esta mezcla; c) Que este beneficio debería ser balanceado con un riesgo mayor de complicaciones de la técnica y un riesgo desconocido de complicaciones tardías o poco frecuentes; d) Que el procedimiento solamente se recomendaría si se realizara en el mismo centro en el cual se realizó el ECA y con la presencia o supervisión directa de uno de los operadores con experiencia en esa técnica.

Desde el punto de las opiniones de cada participante en la junta, para nuestro colega internista, la nueva mezcla no debía recomendarse por el momento, dado que:

a) El paciente tenía alto riesgo de complicaciones posoperatorias sobre las que no hay información disponible con el uso de esta técnica; b) El nuevo procedimiento no redujo significativamente el tiempo quirúrgico y c) Generó un exceso de complicaciones intraoperatorias. Los colegas anestesiólogo y el oftalmólogo coincidieron en su posición a favor de la nueva técnica porque ésta: a) Redujo el tiempo de latencia; b) Produjo un mejor tiempo de analgesia, ventaja que podría incidir en otros desenlaces, tales como una menor proporción de intervenciones, efectos secundarios de analgésicos, costos de atención, etc y c) Debido a la mayor eficiencia que estas ventajas de la nueva intervención podrían generar en centros de alto volumen de cirugía oftálmica. Además, el equipo tenía confianza en que con mayor experiencia las complicaciones intraoperatorias se reducirán cada vez más. De hecho, el equipo ya contaba con una experiencia (no sistemática) de "varios años", en los que probablemente las complicaciones graves -de haber sucedido- serían más fácilmente recordadas. La decisión final se tomaría en esa cita al escuchar las consideraciones del paciente y su entorno.

Como recomendación para plantear dentro del grupo de investigadores se sugirió la extensión del ECA a un número y espectro mayor de pacientes y centros, de tal suerte que se pudiese evaluar la presencia de complicaciones, se redujera la posibilidad de error alfa y se aumentara la validez externa del actual ECA. Igualmente, con un mayor tamaño, el nuevo ECA facilitaría una evaluación de los procedimientos en diferentes niveles de experiencia de los operadores. Finalmente, se recomendó para el nuevo ECA el establecimiento de algunos mecanismos que garanticen el enmascaramiento de la intervención asignada a través del procedimiento y la evaluación de las complicaciones posoperatorias.

Summary

A well-conducted randomized controlled trial (RCT) provides health care givers with good-quality information on the effect of either preventive or therapeutic interventions. Users of RCTs should, therefore, be aware of the sources of RCTs that allow them easy and fast access to reliable information on the questions rose in their own clinical practice. Once accessed, if provided with tools to judge both RCTs validity and applicability, a reader will make a better use of it. By using an RCT recently published in MEDUNAB, this paper discusses a) The reasons why RCTs are sought as primary source to guide clinical decision-making, using the so-called grades of recommendations as standard; b) Some widely and easily available sources to find RCTs using the internet, specifically the Cochrane library, some secondary journals and the PubMed service; c) the criteria used to evaluate RCTs's quality, such as those proposed by the evidence based working group as well as Jadad's scale; d) Some measures to estimate the size of the effect of the health interventions commonly used in the RCTs; e) Finally, some criteria for applying RCTs in the day-to-day clinical practice.

Key words: Controlled trials, methodology.

Referencias

- Villar JC. Elementos para la valoración y el uso práctico de los ensayos clínicos. Parte I: Estructura básica y terminología usual. MEDUNAB 2001;4:107-20.
- Harris WS, Gowda M, Kolb JW, et al. A randomized, controlled trial of the effects of remote, intercessory prayer on outcomes in patients admitted to the coronary care unit. Arch Intern Med 1999;159:2273-8.
- Hoover DR, Margolick JB. Questions on the design and findings of a randomized, controlled trial of the effects of remote, intercessory prayer on outcomes in patients admitted to the coronary care unit. Arch Intern Med 2000;160:1875-6.
- Jadad AR, Rennie D. The randomized controlled trial gets a middle-aged checkup. JAMA 1998;279:319-20.
- Sackett DL. Time to put the Canadian Institutes of Health Research on trial. CMAJ 1999;161:1414-5.
- Oxman AD, Sackett DL, Guyatt GH. Users' guides to the medical literature. I. How to get started. The Evidence-Based Medicine Working Group. JAMA 1993;270:2093-5.
- Rangel GW, Chona JE, García LM. Bloqueo peribulbar adicionando opioides en analgesia posoperatoria para cirugía oftálmica. Un ensayo clínico controlado. MEDUNAB 2001; 4:5-12.
- The Cochrane collaboration. The Cochrane library. The Cochrane library. 2001. <http://www.cochranelibrary.com/clib/>
- Evidence based medicine. EBM on line. BMJ Journals website. 2001 <http://ebm.bmjournals.com/>
- American College of Physicians - American Society of Internal Medicine. ACP-ASIM online. ACP Journal website. 2001. <http://www.acponline.org/journals/acpj/jcmenu.htm>
- National library of medicine. PubMed. PubMed website. 2001. <http://www.ncbi.nlm.nih.gov/PubMed/>
- Tangkanakul C, Counsell C, Warlow C. Local versus general anesthesia for carotid endarterectomy (Cochrane review). In: The Cochrane library, Issue 4, 2.001 Oxford: Update software.
- Hollis L.J, Burton M.J, Millar J.M. Perioperative local anesthesia for reducing pain following tonsillectomy (Cochrane review). In: The Cochrane library, Issue 4, 2.001 Oxford: Update software.
- Sousa Escandon MA, González GC, González Fernández MM. ¿Cuál idioma hablará Medline en el nuevo milenio? Arch Esp Urol 2000;53:93-9.
- Haiqi Z, Yamazaki S, Urata K. The tendency toward English-language papers in Medline. Bull Med Libr Assoc 1997; 85:432-4.
- Jadad AR. Types of randomised controlled trials. In: Randomised controlled trials. London: BMJ books; 1998:11-27.
- García-Valdecasas-Campelo E, González-Reimers E, López-Lirola A, Rodríguez-Rodríguez E, Santolaria-Fernández F. Acute rhabdomyolysis associated with cerivastatin therapy. Arch Intern Med 2001;161:893.
- Wooltorton E. Bayer pulls cerivastatin (Baycol) from market. CMAJ 2001;165:632.
- Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med. 2001;134:663-94.
- Yusuf S, Collins R, Peto R. Why do we need some large, simple randomized trials? Stat Med 1984;3:409-22.
- Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA 1993;270:2598-601.
- Jadad AR. Assessing the quality of RCT's: Why, what how and by whom. In: Randomised controlled trials. London: BMJ books; 1998:46-61.
- Boutitie F, Boissel JP, Connolly SJ, et al. Amiodarone interaction with beta-blockers: analysis of the merged EMIAT (European Myocardial Infarct Amiodarone Trial) and CAMIAT (Canadian Amiodarone Myocardial Infarction Trial) databases. The EMIAT and CAMIAT Investigators. Circulation. 1999;99:2268-75.
- Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996;17:1-12.
- The consort group. The consort statement. The consort group website. 2001. <http://www.consort-statement.org/>

Apéndice 1

Referencias seleccionadas de Medline

- Agrawal K, Saxena RC, Nath R, Saxena S. Local anesthesia by peribulbar block for cataract extraction in an eye relief camp. A double-masked, randomized controlled trial. Online J Curr Clin Trials. 1993;Doc No 40:3069.
- Anders N, Heuermann T, Ruther K, Hartmann C. Clinical and electrophysiologic results after intracameral lidocaine 1% anesthesia: a prospective randomized study. Ophthalmology. 1999;106:1863-68.
- Azuara-Blanco A, Moster MR, Marr BP. Subconjunctival versus peribulbar anesthesia in trabeculectomy: a prospective, randomized study. Ophthalmic Surg Lasers. 1997;28:896-99.
- Bedi A, Carabine U. Peribulbar anaesthesia: a double-blind comparison of three local anaesthetic solutions. Anaesthesia. 1999;54:67-71.
- Costa P, Papurel BG, Coaloa M, Villa C, Ravera E, Hellmann F et al. [Loco-regional block in ophthalmic surgery: single drug or drug combination with hyaluronidase? Randomized prospective study]. Minerva Anestesiol. 1999;65:775-83.
- Davis PL, O'Connor JP. Peribulbar block for cataract surgery: a prospective double-blind study of two local anesthetics. Can J Ophthalmol. 1989;24:155-58.
- Gioia L, Prandi E, Codenotti M, Casati A, Fanelli G, Torri TM et al. Peribulbar anesthesia with either 0.75% ropivacaine or a 2% lidocaine and 0.5% bupivacaine mixture for vitreoretinal surgery: a double-blinded study. Anesth Analg. 1999;89: 739-42.
- Luchetti M, Magni G, Marraro G. A prospective randomized double-blinded controlled study of ropivacaine 0.75% versus bupivacaine 0.5%-mepivacaine 2% for peribulbar anesthesia. Reg Anesth Pain Med. 2000;25:195-200.
- Perello A, George J, Skelton V, Pateman J. A double-blind randomised comparison of ropivacaine 0.5%, bupivacaine 0.375% - lidocaine 1% and ropivacaine 0.5% - lidocaine 1% mixtures for cataract surgery. Anaesthesia. 2000;55:1003-7.