

**Artículo especial****Revisiones sistemáticas con metaanálisis de ensayos clínicos: ¿es evidencia suficiente?****Josep M. Garcia-Alamino<sup>a,\*</sup> y Manuel Lopez-Cano<sup>b</sup>**<sup>a</sup> Programme in Evidence-Based Healthcare, University of Oxford, Oxford, Reino Unido<sup>b</sup> Departamento de Cirugía General, Unidad de Cirugía de la Pared Abdominal y Grupo de Investigación en Cirugía General y del Aparato Digestivo, Institut de Recerca Vall d'Hebron (VHIR), Hospital Universitario Vall d'Hebron, Universidad Autónoma de Barcelona, Barcelona, España**INFORMACIÓN DEL ARTÍCULO****R E S U M E N****Historia del artículo:**

Recibido el 10 de febrero de 2019

Aceptado el 3 de agosto de 2019

On-line el 14 de octubre de 2019

**Palabras clave:**

Metaanálisis

Revisión sistemática

Análisis secuencial

Tamaño de la información

Las revisiones sistemáticas con metaanálisis son pruebas esenciales para resumir de manera fiable la eficacia de las intervenciones sanitarias. Un metaanálisis basado exclusivamente en un pequeño número de ensayos a menudo no será concluyente o puede aportar resultados positivos falsos. La estimación del tamaño óptimo de la información (OIS) puede reducir el riesgo de obtener resultados falsos positivos y determinar si se requieren ensayos clínicos adicionales. El cálculo del OIS tiene similitudes con el realizado para calcular el tamaño de la muestra de un ensayo clínico. El número requerido de participantes para el metaanálisis debe ser al menos tan grande como un solo ensayo con la potencia adecuada. El concepto del OIS proporciona un marco estadístico en el que la información acumulada es convincente para indicar si se necesitan más ensayos clínicos para abordar los efectos de la intervención.

© 2019 AEC. Published by Elsevier España, S.L.U. Todos los derechos reservados.

**Systematic reviews with meta-analysis of clinical trials: Is there enough evidence?****A B S T R A C T****Keywords:**

Meta-analysis

Systematic review

Trial sequential analysis

Optimal information size

Systematic reviews with meta-analysis are essential tools to reliably summarize the effectiveness of health interventions. A meta-analysis based exclusively on a small number of trials will often not be conclusive or may yield false positive results. The estimation of the optimal information size (OIS) can reduce the risk of obtaining false positive results and determine if additional clinical trials are required. The estimation of the OIS is very similar to that used to estimate the sample size of a clinical trial. The required number of participants for the meta-analysis should be at least as large as a single trial with adequate power. The OIS concept provides a statistical framework in which the accumulated information is convincing to indicate whether more clinical trials are needed to address the effects of the intervention.

© 2019 AEC. Published by Elsevier España, S.L.U. All rights reserved.

\* Autor para correspondencia.

Correo electrónico: [Josepmariagarciaa@gmail.com](mailto:Josepmariagarciaa@gmail.com) (J. M. Garcia-Alamino).<https://doi.org/10.1016/j.ciresp.2019.08.006>

0009-739X/© 2019 AEC. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

## Introducción

La revisión sistemática (RS) con metaanálisis es una prueba esencial para sintetizar de manera fiable la eficacia de las intervenciones sanitarias<sup>1</sup>. La RS con metaanálisis (MA) es un trabajo de investigación que identifica el máximo número de estudios, evalúa su calidad y resume sus resultados utilizando una metodología científica rigurosa y explícita. Los cirujanos la utilizan para mantenerse actualizados en su campo y es esencial para desarrollar guías de práctica clínica<sup>2</sup>. Ciertas agencias estatales de financiación de la investigación solicitan conocer las RS existentes antes de financiar un nuevo proyecto de investigación<sup>3</sup>.

La RS puede incluir el MA como un método estadístico para resumir los resultados de los ensayos clínicos individuales<sup>4</sup>. El objetivo principal del MA es proporcionar una estimación más precisa que combine la información numérica de los ensayos clínicos individuales. Sin embargo, un MA basado exclusivamente en un pequeño número de ensayos a menudo no será concluyente<sup>5</sup> o puede dar lugar a errores de tipo I, es decir, aceptar que hay una diferencia entre las 2 intervenciones evaluadas a pesar de que esa diferencia no existe<sup>6</sup>. En la literatura hay diversos trabajos en los que se concluye que los resultados de un MA con un amplio número de pacientes son probablemente más válidos que aquellos con un pequeño número de pacientes y eventos (<200 eventos)<sup>7-9</sup>.

Aunque actualmente se discute e incluso cuestiona si se deben continuar utilizando los valores de significación estadística (valores de  $p$ ) de manera categórica<sup>10</sup>, hasta el momento lo más extendido ha sido fijar un valor de riesgo menor o igual al 5% o menor o igual al 1% de cometer un error de tipo I. En este sentido, al menos, ese mismo valor de riesgo es el que se debería preservar en el MA. Actualizar un MA cada vez que aparece un nuevo ensayo clínico (algo habitual) incrementa el riesgo de sobreestimar el efecto observado en el MA y esto está principalmente relacionado con el efecto de multiplicidad estadística, lo que significa que se puede asumir un beneficio cuando realmente no existe tal beneficio. Es decir, cada vez que se incorpora un nuevo ensayo o se analiza un *outcome* diferente, se irá consumiendo valor de  $\alpha$  y el nivel de error tipo I podrá llegar a ser superior a 5%. Se estaría obteniendo un resultado significativo que, en realidad, correspondería a una sobreestimación del efecto real de la intervención. Sin embargo, es probable que un gran número de clínicos e investigadores sean ajenos a estas situaciones y acepten como «definitivos» los resultados de un MA.

El objetivo de este trabajo es describir el concepto del tamaño muestral en el MA, del tamaño óptimo de la información (*optimal information size*; OIS) de un MA, su cálculo, su interpretación y la influencia del análisis del OIS en la interpretación de la evidencia.

## Un tamaño de muestra en el metaanálisis

Desde un enfoque estadístico clásico (frecuentista), el cálculo del tamaño muestral es un requisito imprescindible cuando se desarrolla el protocolo de un ensayo clínico. Esto se refleja en el número de pacientes u otras unidades experimentales que

deben incluirse en un estudio para responder de manera fiable a la pregunta de investigación<sup>11</sup>. El tamaño de muestra en el MA puede definirse como el número de participantes y eventos necesarios para detectar o rechazar un efecto de intervención asumido *a priori* en un MA. Aquí también se podría aplicar un criterio similar al de los ensayos clínicos individuales, asumiendo que si la muestra es demasiado pequeña no será posible detectar un efecto y que una muestra demasiado grande puede representar una pérdida de tiempo y dinero, e incluso plantear problemas éticos<sup>12</sup>. Esto se explicaría principalmente por el hecho de continuar realizando ensayos clínicos cuando un MA ya hubiera demostrado un resultado a favor o en contra de la intervención de manera sólida y, por ello, se estaría exponiendo a pacientes a una intervención inefectiva o no se estaría ofreciendo a los pacientes una intervención efectiva en el caso de que se hubiera mostrado así en el MA. Entre la comunidad científica y los profesionales asistenciales, se ha podido llegar a extender la idea de que los MA proporcionan un resultado fiable solo por el hecho de ser un MA.

Sin embargo, ¿cuándo puede un MA considerarse una prueba definitiva?, ¿es posible realizar un cálculo del tamaño muestral en el MA al igual que en el ensayo clínico?

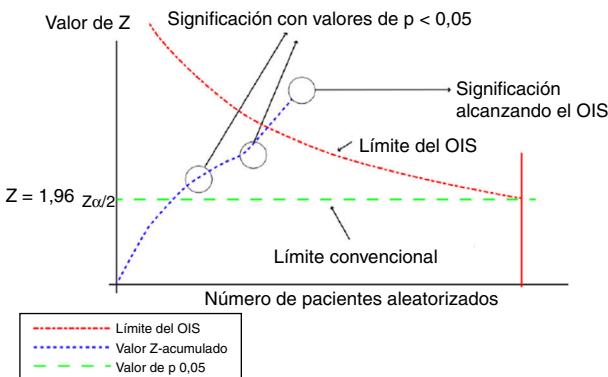
Casi en paralelo a los inicios del movimiento de la medicina basada en la evidencia, el número de MA publicados anualmente se ha incrementado de manera casi exponencial. En la actualidad, se estima que se publican alrededor de 8.000 revisiones sistemáticas al año<sup>13</sup>. Sin embargo, no todos los MA incluyen el número apropiado de pacientes para obtener conclusiones fiables y la falta de un número de pacientes y eventos adecuado es una de las razones de la imprecisión en el efecto estimado<sup>14</sup>.

## Cálculo del tamaño óptimo de la información en el metaanálisis

En un esfuerzo por obtener conclusiones más fiables mediante el MA, la estimación del OIS puede reducir el riesgo de obtener resultados falsos positivos y determinar si se requieren ensayos clínicos adicionales en un área de conocimiento. Concretamente, el OIS puede definirse como la cantidad mínima de información requerida en la literatura para llegar a conclusiones confiables sobre una intervención<sup>15</sup>. El cálculo del OIS tiene similitudes con el realizado para el tamaño de la muestra de un ensayo clínico. El concepto de OIS para el MA fue inicialmente desarrollado y propuesto por Pogue en el año 1998<sup>15</sup>. Para ejecutar el cálculo del OIS se han propuesto varias opciones. El método más sencillo para calcular este número de pacientes, similar al cálculo del tamaño muestral, es ejecutarlo como si se tratara de un ensayo clínico con potencia adecuada. Esto se puede realizar asumiendo un valor de  $\alpha$  (falso positivo o error de tipo I), de  $\beta$  (falso negativo o error de tipo II), una potencia, un tamaño del efecto y una tasa de eventos en el grupo control.

Este análisis se puede representar de forma gráfica, de manera similar a la de los ensayos clínicos secuenciales, según se muestra en la figura 1.

Para hacer este cálculo se puede utilizar un software de cálculo de tamaño de la muestra como, por ejemplo, el de



**Figura 1 – Representación gráfica del análisis secuencial aplicado al metaanálisis. Adaptada con permiso de Imberger et al.<sup>22</sup>.**

En este gráfico la línea del valor Z-acumulado corresponde a la información acumulada a partir de cada uno de los ensayos incluidos en el metaanálisis. En el momento en que la línea del valor Z-acumulado atraviesa la línea de valor  $p < 0,05$  (límite convencional), que corresponde al valor tradicional de significación, se obtiene un resultado estadísticamente significativo, pero ello no indica que se haya alcanzado el OIS. Se alcanzará el volumen adecuado de información cuando la línea de valor Z-acumulado atraviese la línea del límite del OIS, tal y como se muestra en el gráfico.

acceso gratuito PS-Power (<http://biostat.mc.vanderbilt.edu/wiki/Main/PowerSampleSize>). Una década después en el año 2008, Brok et al.<sup>16</sup> propusieron incluir la heterogeneidad estadística en la estimación del OIS considerando que es un parámetro esencial en el MA. La heterogeneidad estadística cuantifica la variabilidad entre los resultados de los estudios y esta variabilidad puede ser debida a las diferencias reales de planteamiento y ejecución entre los estudios incluidos o a otras causas. Para la estimación de este OIS, incluyendo la heterogeneidad, se ha diseñado al menos un software específico y de libre acceso que permite el cálculo del OIS asumiendo un valor de heterogeneidad que puede introducirse manualmente o asumirse por defecto según el cálculo realizado por el software (<http://www.ctu.dk/tsa/downloads.aspx>).

Tal y como se ha mencionado en unos párrafos previos, en la estimación del OIS un paso clave es decidir qué parámetros incluir en su cálculo y un paso no menos importante pero más controvertido es cómo estimar esos parámetros que se asumirán para el cálculo. Con relación al primer punto, los parámetros necesarios pueden ser el efecto estimado, la tasa de eventos en el grupo control y la heterogeneidad ( $I^2$ ). Hay otros parámetros también necesarios que son el valor de  $\alpha$ , el valor de  $\beta$  y la potencia estadística. Para estos 3 últimos sí existe un mayor consenso y suelen asumirse los siguientes valores:  $\alpha = 0,05$ ;  $\beta = 95\%$ ; potencia = 80%. En el caso del efecto estimado, hay autores que utilizan el efecto estimado en el MA, otros utilizan un valor de 20% expresado como reducción relativa del riesgo. Acerca de la tasa de eventos, principalmente la tasa de eventos en el grupo de control, algún autor

propone que la opción más estable es el valor de la mediana de la proporción de eventos en el grupo control<sup>17</sup>. Respecto a la heterogeneidad, se suele tomar el valor de la heterogeneidad observada en el MA. Wetterslev et al. han propuesto lo que han denominado diversity, expresado como  $D^2$ , y según los autores parece una mejor alternativa que  $I^2$  para considerar la variación del modelo en cualquier MA de efectos aleatorios<sup>18</sup>. Aunque existen otras opciones quizás más apropiadas, como el uso de valores estimados con distribuciones *a priori* de la heterogeneidad<sup>19,20</sup>.

## Interpretación del tamaño óptimo de la información

En la figura 2 se muestran 3 escenarios diferentes del cálculo del OIS y su interpretación que perfectamente podrían ser aplicables a MA en el ámbito de la cirugía.

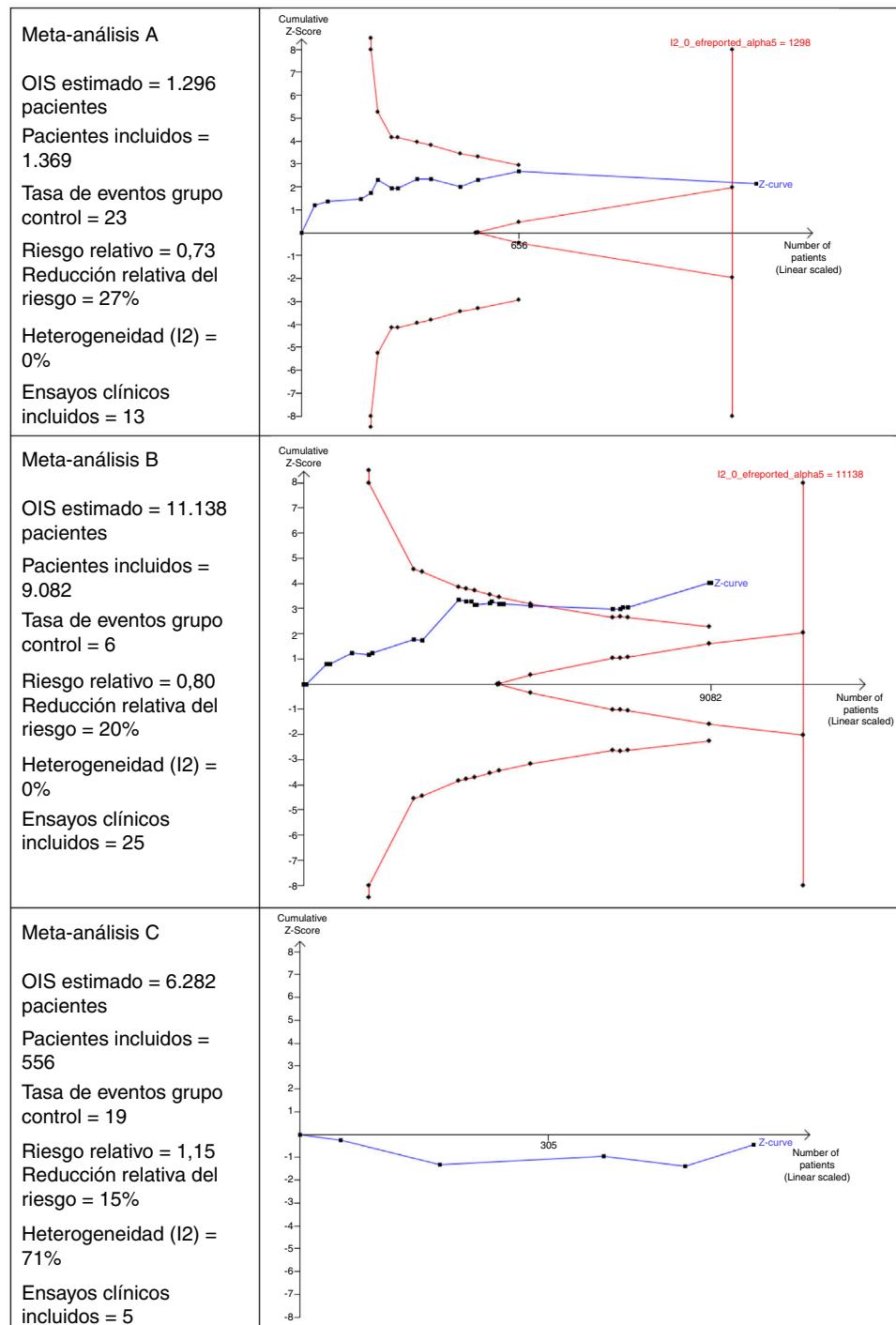
En el MA A se observa que incluye 13 ensayos clínicos con un total de 1.369 pacientes. Este número de pacientes es suficiente para alcanzar el OIS de 1.298 pacientes y, por ello, la curva z cruza la línea vertical. Este es un buen ejemplo en el que se alcanza un tamaño muestral suficiente para tener confianza en el resultado obtenido en el MA.

En el MA B se puede observar que hay 25 ensayos clínicos con un total de 9.082 pacientes. Este número es insuficiente para alcanzar el OIS de 11.138 pacientes y, por ello, la curva z no cruza la línea vertical, aunque está cerca de alcanzarla porque, además, se observa la tendencia y cruza la línea de la curva de significación (curva descendente del límite de pacientes que deben incluirse).

En el MA C se puede observar que incluye un total de 5 ensayos clínicos que acumulan 556 pacientes. El OIS estimado es de 6.282 pacientes. Debido a este bajo número de pacientes acumulados, ni siquiera es posible dibujar la curva descendente. Otro detalle interesante de esta figura es que la curva z transcurre por debajo de la línea horizontal debido a que el riesgo relativo observado es  $>1$ .

## Tamaño óptimo de la información e interpretación de la evidencia

El concepto de OIS proporciona un marco estadístico en el que la información acumulada aporta fiabilidad para indicar si se necesitan más ensayos clínicos que evalúen los efectos de la intervención en la población general. El cálculo del OIS es una herramienta estadística que permite tener información más precisa sobre la estimación del efecto observado en un MA. También puede informar de si el volumen de información acumulada en el MA, principalmente el número de estudios, bajo ciertas asunciones, es suficiente para responder a la pregunta de investigación. La estimación del OIS ya ha sido considerada como útil por el GRADE working group, que evalúa la calidad de la evidencia y establece el grado de recomendación para una determinada intervención médica o quirúrgica. El sistema GRADE considera esta estimación del OIS un parámetro informativo para evaluar la imprecisión de los resultados obtenidos en un MA. De esta manera, también determina la confianza que aporta el resultado y, con base en



**Figura 2 – Ejemplos de cálculo del tamaño óptimo de la información en el que se observan 3 escenarios diferentes (metaanálisis A, B y C).**

ello, establece un grado de recomendación u otro para la intervención estudiada<sup>14</sup>.

En las revisiones sistemáticas con MA se está haciendo un uso creciente del OIS: ya existen varios MA publicados en el ámbito quirúrgico que incorporan el cálculo del OIS<sup>21-23</sup>. A pesar de ello, aún está en desarrollo y existe debate sobre cómo debe realizarse la estimación de los parámetros

recomendados para el cálculo<sup>24</sup>. No obstante, y a pesar de sus limitaciones metodológicas, se hace necesario disponer de una herramienta de este tipo que nos permita tener una información más precisa para interpretar la evidencia y evite en más de una ocasión la conocida frase *no hay suficiente evidencia para evaluar el efecto de esta intervención*. Son necesarios más estudios para investigar su efecto. En el caso de que se deba

utilizar esta frase, se podrá indicar con más precisión qué número de pacientes es necesario en dichos estudios para disponer de una información más precisa y fiable.

## Financiación

La presente investigación no ha recibido ninguna beca específica de agencias de los sectores público, comercial ni sin ánimo de lucro.

## Conflictos de intereses

Ninguno.

## B I B L I O G R A F Í A

1. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*. 2009;339:b2700.
2. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med*. 2009;6:e1000097. 10.1371/journal.pmed.1000097.
3. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Reprint-preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PhysTher*. 2009;89:873-80.
4. Glass GV. Primary, secondary and meta-analysis of research. *Educ Res*. 1976;5:3-8.
5. Davey SG, Egger M. Meta-analysis. Unresolved issues and future developments. *BMJ*. 1998;316:221-5.
6. Hu M, Cappelleri JC, Lan.F K.K.. Applying the law of iterated logarithm to control type I error in cumulative meta-analysis of binary outcomes. *Clin Trials*. 2007;4:329-40.
7. Pereira TV, Ioannidis JP. Statistically significant meta-analyses of clinical trials have modest credibility and inflated effects. *J Clin Epidemiol*. 2011;64:1060-9.
8. Thorlund K, Imberger G, Walsh M, Chu R, Gluud C, Wetterslev J, et al. The number of patients and events required to limit the risk of overestimation of intervention effects in meta-analysis: A simulation study. *PLoS One*. 2011;6:e25491. 10.1371/journal.pone.0025491.
9. Imberger G, Thorlund K, Gluud C, Wetterslev J. False-positive findings in Cochrane meta-analyses with and without application of trial sequential analysis: An empirical review. *BMJ Open*. 2016;6:e011890. 10.1136/bmjopen-2016-011890.
10. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature*. 2019;567:305-7.
11. Noordzij M, Dekker FW, Zoccali C, Jager KJ. Sample Size Calculations. *Nephron Clin Pract*. 2011;118:c319-23.
12. Emerson SS, Kittelson JM, Gillen DL. Frequentist evaluation of group sequential clinical trial designs. *Stat Med*. 2007;26:5047-80.
13. Ioannidis JP. The mass production of redundant misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q*. 2016;94:485-514.
14. Guyatt GH(1), Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6: Rating the quality of evidence-imprecision. *J Clin Epidemiol*. 2011;64:1283-93.
15. Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomised controlled trials. *Lancet*. 1998;351:47-52.
16. Brok J, Thorlund K, Gluud C, Wetterslev J. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *J Clin Epidemiol*. 2008;61:763-9.
17. Hayden J, Tomlinson G, Ni A, Parker R, Cartwright J, van Tulder M, et al. Approaches to estimate and present baseline risks considerations for Cochrane reviews evidence-informed practice integration of: 19th Cochrane Colloquium Madrid. 2011.
18. Wetterslev J, Thorlund K, Brok J, Gluud C. Estimating required information size by quantifying diversity in random-effects model meta-analyses. *BMC Med Res Methodol*. 2009;9:86. 10.1186/1471-2288-9-86.
19. Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JP. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. *Int J Epidemiol*. 2012;41:818-27.
20. Garcia-Alamino JM, Bankhead C, Heneghan C, Piddock N, Perera R. Impact of heterogeneity and effect size on the estimation of the optimal information size: Analysis of recently published meta-analyses. *BMJ Open*. 2017;7:e015888.
21. López-Cano M, Martín-Domínguez LA, Pereira JA, Armengol-Carrasco M, García-Alamino JM. Balancing mesh-related complications and benefits in primary ventral and incisional hernia surgery. A meta-analysis and trial sequential analysis. *PLoS One*. 2018;13:e0197813. 10.1371/journal.pone.0197813.
22. Antoniou SA, García-Alamino JM, Hajibandeh S, Hajibandeh S, Weitzendorfer M, Muysoms FE, et al. Single-incision surgery trocar-site hernia: An updated systematic review meta-analysis with trial sequential analysis by the Minimally Invasive Surgery Synthesis of Interventions Outcomes Network (MISSION). *Surg Endosc*. 2018;32:14-23.
23. López-Cano M, Brandsma HT, Bury K, Hansson B, Kyle-Leinhase I, Alamino JG, et al. Prophylactic mesh to prevent parastomal hernia after end colostomy: A meta-analysis and trial sequential analysis. *Hernia*. 2017;21:177-89.
24. Garcia-Alamino JM. Assessing the role of optimal information size in systematic reviews [tesis doctoral]. Oxford: University of Oxford; 2018.