

## SUPPLEMENTAL INFORMATION 2

### SPANISH TRANSLATION OF THE CHECKLIST FOR CRITICAL APPRAISAL OF INDIRECT COMPARISONS

#### PARTE A/ ¿Son válidos los resultados de la comparación indirecta?

<p>- <b>¿Está justificado realizar una comparación indirecta (CI)?</b></p> <p><i>¿Existe una comparación directa? En general, se considera que las comparaciones directas tienen menos sesgos que las indirectas. Si hay una comparación directa y una indirecta puede ser interesante valorar si llegan a la misma conclusión. Si no hay diferencias aumentará nuestra fiabilidad en el resultado y si hay diferencias debemos cuestionarnos a qué pueden deberse esas diferencias.</i></p> <p><i>Para valorarse si se puede hacer una comparación indirecta se debe cumplir:</i></p> <ul style="list-style-type: none"><li>- <i>¿Están claros el objetivo, la población, los métodos (incluidas asunciones) y los resultados? (si no hay transparencia en todos estos aspectos no se puede o dificulta la evaluación de la CI)</i></li><li>- <i>¿Los estudios comparados son comparables clínicamente? (Características de la población estudiada, condiciones coexistentes, cointervenciones, dosificación, duración del tratamiento y resultados evaluados que influyen o modifican la magnitud del efecto de la intervención)</i></li><li>- <i>¿La medida de resultados es comparable en cuanto a su definición y su presentación?</i></li><li>- <i>¿La medida de resultados es clínicamente relevante?</i></li></ul>	<p><b>SÍ      NO</b> <b>NO SÉ</b></p>
<p>- <b>¿El método de CI es apropiado?</b></p> <ul style="list-style-type: none"><li>- <i>Las comparaciones indirectas no ajustadas (que no comparan el efecto relativo, sino que cogen un brazo de un estudio sin considerar el comparador) se consideran no apropiadas.</i></li><li>- <i>Métodos de comparaciones indirectas: Bucher, Metanálisis de redes de tratamiento, mezclas de comparaciones indirectas, metanálisis de múltiples tratamientos.</i></li></ul>	<p><b>SÍ      NO</b> <b>NO SÉ</b></p>
<p>- <b>¿Se realizó una búsqueda adecuada de todos los estudios pertinentes?</b></p> <p><i>¿Se dirige adecuadamente la pregunta para establecer la CI? (Valorar acrónimo PICO [Patient, Intervention; Comparison; Outcome])</i></p> <p><i>¿Están incluidos todos los estudios relevantes?</i></p> <p><i>¿El comparador es relevante?</i></p>	<p><b>SÍ      NO</b> <b>NO SÉ</b></p>

Tras las respuestas a estas tres preguntas, ¿consideras que debes continuar analizando este estudio?.

**SÍ**  
↓  
CONTINUAR

**NO**  
↓  
PARAR

**PARTE B/ PREGUNTAS DE DETALLE: LISTA DE COMPROBACIÓN PARA LA EVALUACIÓN DE UNA COMPARACIÓN INDIRECTA**

ITEM	OPCIONES	FIABILIDAD	COMENTARIO
1 <b>COMPARADOR COMÚN</b> ¿Los comparadores y los estudios están representados gráficamente? ¿Cuántos comparadores intermedios existen entre dos tratamientos a comparar A y B que se pretenden comparar indirectamente?	1	ALTA	Con más de 2 comparadores entre A y B, los intervalos de confianza (IC) de la diferencia estimada indirectamente son muy amplios y existe mucho riesgo al interpretar los resultados si no se consideran en todo momento los IC. La elaboración de un gráfico con todo los estudios incluidos, proporciona una valiosa información para la interpretación de la evidencia disponible
	2	DUDOSA	
	más de 2	BAJA	
2 <b>ESTRATEGIA DE BÚSQUEDA</b> ¿Se utilizaron métodos sistemáticos para localizar todos los estudios relevantes? (incluyeron búsqueda en diferentes bases de datos así como la estrategia de búsqueda, siguieron las referencias, no usaron límites estrictos, valoraron sesgo de publicación, etc)	Si	ALTA	Para intentar localizar los artículos relevantes debería abordarse la búsqueda en diferentes bases de datos. Debe especificarse la estrategia para que pueda reproducirse y comprobar que cumple los criterios especificados.
	Los posibles sesgos y sus efectos son analizados	ACEPTABLE	
	No se cumple ni se indica el sesgo	BAJA	
3 <b>CRITERIOS DE SELECCIÓN DE ESTUDIOS PARA LA COMPARACIÓN</b> ¿Están definidos previamente? ¿Son razonables? ¿Se han aplicado de forma sistemática?	Si, cumple todo	ALTA	Una selección sesgada de estudios puede condicionar totalmente los resultados, por lo que los criterios de selección deben ser explícitos y estar definidos antes de la selección. Es importante poder comprobar que no hay una selección sesgada de los estudios, que se detalle si se han excluido estudios, el motivo, y este debe ser razonable. Si se ha realizado una revisión sistemática debería incluir un diagrama de flujo con la información localizada y seleccionada
	Alguno de estos puntos es mejorable, aunque aceptable	ACEPTABLE	
	No se cumple alguno de los puntos o se hace de forma inaceptable (sesgo importante)	BAJA	
4 <b>VARIABLE COMÚN</b> ¿Se utiliza en los estudios la misma variable definida de la misma forma (incluyendo mismo horizonte temporal) y presentada de la misma manera?	Si	ALTA	Comparar distintas variables o medidas en tiempos distintos, puede condicionar totalmente los resultados.
	No, pero la diferencia entre ambas variables es muy poco relevante. Se especifican y discuten las diferencias.	DUDOSA	
	No	BAJA	
5 <b>RELEVANCIA DE LA VARIABLE</b> La variable elegida para la comparación, ¿es clínicamente la más relevante posible? ¿es la variable principal de los estudios?	Si	ALTA	La elección de una variable poco relevante clínicamente cuando había otras opciones posibles constituye un sesgo inaceptable. En cambio, si la elección de una variable relevante, coincide con que sea la principal de los estudios, aporta una mayor validez interna. La selección de la variable debe estar justificada y la justificación ser razonable.
	Es la más relevante, pero no la principal de los estudios	ACEPTABLE	
	No es la más relevante posible, pero es razonable	DUDOSA	
6 <b>SIMILITUD CLÍNICA, SIMILITUD DE PACIENTES INCLUIDOS</b> Mirar los estudios que se combinan: ¿los criterios de inclusión y exclusión son similares y razonables? Los grupos aleatorizados dentro de cada estudio: ¿eran comparables?	Debería haberse elegido otra variable más relevante clínicamente, o bien no se disponía de ninguna variable medida con relevancia clínica suficiente	BAJA	Es fundamental combinar poblaciones que sean, al menos en los factores que pueden modificar el efecto, similares. Se puede asumir, si es razonable que las diferencias no afectan ni al efecto relativo ni al riesgo basal. En caso de que afectaran al riesgo basal se podría intentar hacer análisis de subgrupos por riesgo basal, y si afecta al efecto relativo hacer análisis de subgrupos basado en variables que modifican el efecto, solo si clínicamente está justificado. Si las variables son continuas se podría hacer metaregresión si hay datos suficientes y está clínicamente justificado.
	Si, con diferencias muy poco relevantes	ALTA	
	Existen diferencias, pero la mayor parte de los pacientes son similares en los diversos estudios. Se compara un subgrupo y los pacientes de los subgrupos que se comparan son similares.	ACEPTABLE	
7 <b>COMPARADOR COMUN</b> El comparador común, ¿está usado en la misma posología y con la misma duración del tratamiento, y se permiten los mismos tratamientos concomitantes que pueden afectar al efecto?	Existen muchas diferencias de población o en los subgrupos comparados entre los estudios.	BAJA	La combinación de estudios con comparadores comunes que no son realmente iguales puede comprometer totalmente la validez de la comparación, si hay dudas que el efecto sea diferente habría que considerarlas como alternativas diferentes. Cuando no hay un comparador común, podrían considerarse fármacos equivalentes (demostrado en un estudio de equivalencia o un estudio de no inferioridad) pero el riesgo de sesgo es mayor.
	Si	ALTA	
	Existen diferencias con muy poca influencia potencial en el resultado	ACEPTABLE	
	Existen diferencias relevantes	BAJA	

8	<b>RESULTADO CUANTITATIVO DEL GRUPO CONTROL</b>  ¿Es similar en los estudios que se combinan en la CI?	Sí. O no existen diferencias, o claramente carecen de relevancia clínica.	ALTA	Para poder combinar los estudios, el resultado del grupo control debería ser similar. No obstante, habría que buscar la causa y si las poblaciones son comparables, una diferencia en el grupo control puede paliarse por un análisis ajustado por el resultado del control (para ampliar información consultar ICWG Report <a href="http://www.health.gov.au/internet/main/publishing.nsf/Content/B11E8EF19B358E39CA25754B000A9C07/\$File/ICWG%20Report%20FINAL2.pdf">http://www.health.gov.au/internet/main/publishing.nsf/Content/B11E8EF19B358E39CA25754B000A9C07/\$File/ICWG%20Report%20FINAL2.pdf</a> ) o un análisis de subgrupos como se ha indicado en la pregunta 6.
		No. Existen diferencias que podrían ser clínicamente relevantes, pero resultan aceptables	ACEPTABLE	
		No. Existen diferencias que ponen en evidencia que las poblaciones y/o la intervención y/o la medida de resultados probablemente presentan diferencias importantes.	BAJA	
9	<b>CALIDAD DE LOS ESTUDIOS COMBINADOS</b>  ¿Cuál es la calidad/validez interna de los estudios combinados? ¿Se dan los resultados de los estudios individuales (la transparencia es imprescindible para valorar la aplicabilidad)?	Los estudios que se combinan son ensayos clínicos bien diseñados.	ALTA	Los estudios observacionales tienen <i>per se</i> , una elevada probabilidad de sesgos. Combinarlos, además, con un comparador común, los multiplica y condiciona una baja validez interna.  La transparencia en los estudios de partida es imprescindible para valoración de la comparación indirecta (se deberían dar no solo los valores relativos sino también los datos reales).  Los estudios observacionales pueden ser relevantes para evaluar la efectividad y seguridad.
		Los estudios que se combinan incluyen algún estudio de cohortes con baja probabilidad de sesgos. Pero existe al menos un ensayo clínico bien diseñado de cada fármaco que se compara, y el peso de los estudios de cohortes (en nº de pacientes) no supera al de los ensayos clínicos para cada fármaco.	DUDOSA	
		No se puede incluir en ninguno de los dos perfiles anteriores	BAJA	
10	<b>TRATAMIENTO DE LAS PÉRDIDAS</b>  En los estudios que se han combinado, ¿las pérdidas son similares? ¿se trataban de la misma manera? ¿el tipo de análisis es el mismo (por protocolo, por intención de tratar-ITT-)	Sí, cumple todo	ALTA	Un tratamiento distinto de los pacientes que abandonan el estudio, si éstos son numerosos, puede condicionar en gran medida los resultados. Puede paliarse parcialmente mediante el análisis por intención de tratar y sería conveniente realizar un análisis de sensibilidad con los distintos escenarios
		Existen diferencias pero no son relevantes	ACEPTABLE	
		Falta información, o bien existen diferencias relevantes en alguno de estos ítems, pero no anulan completamente la validez de la comparación. Sin embargo, el análisis es ITT en todos los estudios que se combinan.	DUDOSA	
11	<b>FIABILIDAD DE LOS DATOS DE PARTIDA</b> ¿Se han extraído los datos de los estudios? ¿Se han sometido a algún tipo de extrapolación/manipulación previa?	No se puede incluir en ninguno de los dos perfiles anteriores	BAJA	La manipulación o extrapolación de datos supone ya una merma de validez que además se multiplica por la comparación indirecta.
		Los datos coinciden con los publicados	ALTA	
12	<b>AJUSTE POR GRUPO CONTROL</b> ¿Se trata de una comparación ajustada por el grupo control (método de Bucher, Network meta-análisis, etc.), es decir se comparan los efectos relativos respecto al grupo control de cada estudio?	No coinciden, se han manipulado o extrapolado previamente	BAJA	La comparación se tiene que hacer entre los efectos relativos respecto al grupo control (comparador común) en cada estudio.
		Sí	ALTA	
13	<b>ANÁLISIS DE DATOS/ESTADÍSTICO EN UNA CI O EN UN NMA</b> El análisis ¿es adecuado? ¿son razonables las asunciones realizadas? Si es necesario el análisis de sensibilidad ¿es adecuado?	No	BAJA	Las simulaciones a menudo se basan en asunciones subjetivas, y su validez interna es menor. La superposición de intervalos de confianza es una práctica errónea. En los metanálisis en red son preferibles los métodos bayesianos. Si se utilizan se debe describir los valores a priori utilizados, el número de iteraciones, la convergencia y ser estos razonables. Si son necesarias muchas iteraciones para conseguir la convergencia puede ser un indicador de heterogeneidad. En el metanálisis en red si se usa un modelado de efectos fijos se tiene que comprobar que no existe heterogeneidad, en general es mejor utilizar el modelado de efectos aleatorios. Se comprobará la heterogeneidad, si existe se debe analizar la causa y corregir si es posible (ver pregunta 6). Si existe inconsistencia (ver pregunta 15) probablemente se deba a que existe heterogeneidad. Debería existir una descripción de los diferentes escenarios en el análisis de sensibilidad. Valorar el ajuste del modelo.
		Sí. Cuando solo hay 2 alternativas comparadas vs un comparador común con un ensayo cada una, el método de Bucher es adecuado. En el resto de los casos es adecuado el Metanálisis en red. Los métodos deben estar claramente descritos ser adecuados y las asunciones ser razonables.	ALTA	
14	<b>SIMILITUD DE LOS ESTUDIOS</b> ¿Se detecta alguna otra característica diferencial de los estudios que se combinan, que haga inaceptable la comparación de los resultados?	No	ALTA	Puede haber otras causas de heterogeneidad específicas, no incluidas en ítems anteriores, que dificulten la comparabilidad de los distintos estudios utilizados.
		Se han detectado otras diferencias entre los estudios, pero no condicionan los datos de forma muy relevante	DUDOSA	
		Se ha detectado alguna diferencia que condiciona ampliamente los datos y hace incomparables los resultados de los estudios.	BAJA	

15	<p>Contestar sólo si están publicadas comparaciones directas</p> <p><b>¿EXISTE INCONSISTENCIA?</b></p> <p>¿Se observan diferencias en las estimaciones de los efectos de los tratamientos entre las comparaciones directas e indirectas?</p>	<p>No. Se ha testado la inconsistencia y no es significativa. Existen métodos para testar la inconsistencia como la diferencia estandarizada o mediante métodos bayesianos como el modelo de inconsistencia o el node splitting.</p>	ALTA	<p>Debe distinguirse claramente si se observan diferencias significativas entre la comparación directa e indirecta.</p> <p>Si hay inconsistencia es importante estudiar la causa (características de los pacientes, características de los ensayos, otros factores..).</p> <p>Si la comparación indirecta está sesgada es porque lo están las comparaciones directas de las que parte o porque no hay similitud clínica entre los estudios combinados.</p> <p>En general, se consideran menos sesgadas a las comparaciones directas que las indirectas.</p> <p>Para corregir la inconsistencia se puede hacer, en ocasiones, análisis de subgrupos o metaregresión, pero es importante que tenga sentido clínico y que haya datos suficientes para hacerlo. Si no se puede corregir la inconsistencia, el estudio no sirve para la decisión.</p> <p>Si se trata de un metaanálisis en red debe valorarse inconsistencia y todos los aspectos relacionados con la variación en los efectos del tratamiento, incluyendo heterogeneidad dentro de los estudios y entre los estudios</p>
		<p>Si hay diferencias o no se analiza.</p>	BAJA	
16	<p><b>APLICABILIDAD Y VALIDEZ EXTERNA</b></p> <p>¿Las conclusiones se deducen de los resultados?</p> <p>¿Se pueden generalizar estos resultados a nuestra población?</p> <p>¿Incluye la alternativa más relevante para nosotros?</p>	<p>Si</p>	ALTA	<p>Deben considerarse las posibles diferencias con la población a la que queremos aplicar estos resultados y considerar la extrapolación que estamos dispuestos a admitir</p>
		<p>No</p>	BAJA	
17	<p><b>CONFLICTO DE INTERESES</b></p> <p>¿Existe declaración de conflicto de intereses?</p>	<p>Está declarado el conflicto de intereses y no se observa manipulación</p>	ALTA	<p>Una comparación indirecta está sometida a un riesgo de sesgos y manipulaciones, se debe valorar si el conflicto de intereses ha influido en la comparación</p>
		<p>Falta declaración de intereses o se observa manipulación ligada al conflicto de intereses.</p>	BAJA	