

***The Lancet* y el anunciado fin de la homeopatía: revisión crítica de la publicación de Shang et al (2005) y los artículos relacionados que le siguieron**

José Eizayaga

Resumen

En agosto de 2005 *The Lancet* publicó un metanálisis comparativo de ensayos clínicos de homeopatía y ensayos clínicos convencionales que, de acuerdo a sus autores, llega a la conclusión de que la homeopatía es placebo. Este tuvo un impacto muy elevado en los medios y la bibliografía médica. El objetivo de este trabajo es hacer una crítica de esta publicación y una revisión de las publicaciones relacionadas que le siguieron. Conclusiones: la investigación de Shang et al adolece de serios sesgos de selección de los trabajos, sesgos de análisis y probablemente de análisis post hoc que invalidan sus conclusiones. El trasfondo político podría haber tenido una influencia sustancial en los resultados. Los ensayos clínicos homeopáticos son de mejor calidad metodológica que los convencionales y la mayoría de ellos dan un resultado terapéutico positivo.

Palabras clave

Homeopatía; Metanálisis; Ensayos clínicos randomizados

***The Lancet* and the announced end of homeopathy: critical review of Shang et al's (2005) study and subsequent related articles**

Abstract

In August 2005, *The Lancet* published a meta-analysis comparing homeopathic and conventional clinical studies that according to its authors showed homeopathy to be a placebo effect. That article exerted strong impact on the overall media and the medical literature. The aim of the present study was to perform a critical appraisal of that meta-analysis, and a review of the subsequent related publications. The results show that the study by Shang et al exhibits severe selection bias, analysis bias, and possibly also *post hoc* analysis bias, which make its conclusions invalid. Political background factors might have exerted substantial influence on the results. The analyzed homeopathic clinical studies exhibited better methodological quality compared to the conventional ones, and most had positive therapeutic results.

Keywords

Homeopathy; Meta-analysis; Randomized clinical studies

Introducción

En agosto de 2005 *The Lancet* publicó un metanálisis de ensayos clínicos homeopáticos que arrojó como conclusión final que la homeopatía es placebo [1]. La repercusión que este artículo tuvo en estos años, tanto en los medios de difusión pública como en la literatura médica, ha sido sustancial y ha llegado a convertirse en la cita obligada de todos los críticos de la homeopatía, quienes consideran sus conclusiones como una prueba sólida y definitiva en contra de su “poco plausible” efectividad clínica. En la nota editorial intitulada “El fin de la homeopatía”, *The Lancet* festejó los hallazgos del trabajo, llamando la atención al hecho de que a pesar de “150 años de hallazgos desfavorables” y “evidencia cada vez más diluida” la homeopatía con sus “absurdas diluciones [...] parece tener popularidad creciente”. Afirmando que ya ha “pasado el tiempo de los informes sesgados y la inversión en investigación para perpetuar el debate homeopatía versus alopatía”, “los médicos tienen ahora que ser honestos con sus pacientes acerca de la falta de beneficios de la homeopatía” [2].

Las respuestas críticas al trabajo de Shang et al no se hicieron esperar, detallando a su vez una serie de sesgos en los métodos y conclusiones así como otras curiosidades [3-20].

El trabajo de Shang et al es, o al menos así lo pretende de acuerdo al título, un estudio comparativo de ensayos clínicos controlados contra placebo homeopáticos y convencionales. El propósito de este trabajo es analizar y discutir en detalle la publicación de Shang et al y las respuestas críticas que lo sucedieron.

El entorno

Un prejuicio como punto de partida: la homeopatía no puede funcionar

Shang et al parten de la premisa de que si bien el uso de la homeopatía está muy extendido, sus efectos específicos no son plausibles y que a pesar de que la investigación clínica muestra resultados positivos a su favor, la homeopatía **no puede funcionar**. Entonces, como **tiene que haber otros motivos** para estos resultados positivos, se proponen investigar si estos podrían ser explicados por la presencia de los llamados sesgos de investigación y/o sesgos de publicación.

Una de las formas posibles de sesgo son las ideas *a priori* o prejuicios. Por ejemplo, creer fuertemente antes de la investigación que un determinado tratamiento va a actuar de determinada manera. Es claro que es muy difícil librarse completamente de ellos. Pero en Shang et al, el prejuicio es explícito: la homeopatía **no puede funcionar**.

¿De qué manera un prejuicio fuerte y explícito podría incidir en un trabajo de revisión como el de Shang et al? Por ejemplo, influyendo en la selección de los trabajos a ser estudiados y en la forma de ser analizados los resultados. El prejuicio puede conducir a realizar análisis que no han sido establecidos de antemano o a analizar subgrupos de estudios de diversas maneras con el fin de obtener el resultado esperado [4, 9].

Sesgo y calidad metodológica de la investigación

El sesgo o error sistemático (*bias* en inglés) se define como “cualquier tendencia en la recolección, análisis, interpretación, publicación o revisión de los datos que puede dar lugar a conclusiones que difieren sistemáticamente de la verdad” [21]. Muchos tipos de sesgos han sido descritos y entre las mejores herramientas epidemiológicas de que se dispone para evitarlos, se encuentran los llamados ensayos clínicos randomizados y controlados (RCT, de *randomized controlled trials*, ver *Glosario*).

Lo peculiar de los sesgos es que actúan distorsionando los resultados *en un mismo sentido*, ya sea mostrando resultado positivo donde en realidad no lo hay o a la inversa. Un concepto fundamental es que la calidad metodológica de un trabajo de investigación minimiza la posibilidad de que se introduzcan sesgos en él. Esto está señalado por Shang et al, quienes proponen comparar los resultados de los RCT homeopáticos y convencionales de mejor calidad metodológica, que son los que con menor probabilidad están afectados por las diversas formas de sesgo.

Dos hipótesis enunciadas en el trabajo de Shang et al (en la discusión del escrito) son por un lado que los efectos observados en los RCT homeopáticos podrían ser explicados por una combinación de deficiencias metodológicas y sesgos en los reportes, y por otro, que estos mismos sesgos no explicarían los resultados de los RCT convencionales. Dicho de otra manera, que comparando los RCT homeopáticos versus los RCT convencionales de mejor calidad metodológica y por ende con menos chances de padecer sesgos, el tratamiento convencional se mostrará como efectivo y la homeopatía se demostrará como inefectiva.

Otra hipótesis de Shang et al es que los RCT grandes suelen tener mejor calidad metodológica que los pequeños y que los efectos clínicos evidenciados en los metanálisis de RCT homeopáticos podrían estar influenciados por los resultados de los RCT pequeños y de menor calidad. No todos los investigadores comparten este criterio (ver abajo, “El tamaño también importa”).

El estudio de Shang et al es novedoso

La novedad del metanálisis de Shang et al es la comparación de un grupo de RCT homeopáticos con igual número de RCT de tratamiento convencional, emparejados para las mismas enfermedades y variables analizadas. Los metanálisis publicados hasta el momento sólo habían evaluado RCT homeopáticos. El primero y más importante de ellos, publicado también en *The Lancet* 8 años antes, había arrojado un resultado positivo para la homeopatía [22].

Metanálisis

Un metanálisis “consiste en combinar los resultados de estudios individuales, dado que los trabajos son lo suficientemente parecidos como para combinarse” [23]. Al hacerlo se asume que la combinación de los resultados proporcionados por los estudios individuales arroja como resultado una mayor aproximación al efecto terapéutico real de una determinada intervención.

La realización de un metanálisis exige cumplir con una serie de pasos similar a la de un ensayo clínico. La diferencia reside en que en lugar de tratar con pacientes o casos, trata con estudios o ensayos clínicos. Se necesitan objetivos e hipótesis, determinar cómo y dónde serán buscados los estudios, qué criterios de inclusión y exclusión se utilizarán para seleccionarlos, cómo se extraerán los datos de los estudios y finalmente cómo se resumirán y analizarán los datos extraídos. De acuerdo al *Cochrane Handbook for Systematic Reviews*, “sólo pueden obtenerse conclusiones confiables de análisis que fueron realmente especificados antes de inspeccionar los ensayos clínicos” [24]. Este concepto es vital en la investigación: las hipótesis y los métodos deben quedar claramente establecidos antes de la obtención de los datos. El camino opuesto invalida las conclusiones, ya que una vez conocidos los datos es muy tentador formular las hipótesis y los métodos de análisis que favorezcan la obtención de los resultados que se desean obtener.

Un elemento fundamental de los metanálisis es que los estudios que se combinan deben ser comparables entre sí, y una manera de evaluar esto es hacer lo que se llaman cálculos de homogeneidad (ver abajo, “El problema de la heterogeneidad”).

Al combinar los resultados de los trabajos individuales también se asume que, por cuestión de mero azar, los estudios grandes se aproximan más a la magnitud del efecto real de la intervención que los estudios chicos, o dicho de otra manera, “cada estudio individual participa en la magnitud del efecto en relación inversa con su tamaño” [23] (ver abajo, “El tamaño también importa”).

Linde & Jonas, autores del primer metanálisis que analizó RCT homeopáticos mencionado arriba, cuestionan a Shang *et al* por no seguir las recomendaciones para la publicación de metanálisis de QUORUM (publicadas en *The Lancet* en 1999) ni las de *Cochrane Collaboration* y se lamentan que así como su publicación de 1997 había sido mal empleada por los homeópatas como evidencia de que su terapéutica ya estaba probada, ahora “se sienten decepcionados de que una revista médica mayor utilice mal un estudio similar de manera completamente acrítica y polémica”. Y sentencian: “una filosofía subversiva que no sirve ni a la ciencia ni a los pacientes” [3].

Shang *et al* se defenderán de esta crítica afirmando que su trabajo “no es un metanálisis estándar” [25]. Curiosa respuesta, ya que si no se trata de un metanálisis estándar, ¿de qué se trata entonces? ¿Por qué es presentado como si fuera un metanálisis y cuál es su validez?

La calidad metodológica de los RCT importa. E importa mucho

Existe una conocida relación inversa entre la calidad metodológica de un RCT y la magnitud del efecto terapéutico encontrado: cuanto peor la calidad, mejor resultado aparenta tener el tratamiento [26]. Esto también ocurre en los RCT homeopáticos [21]. La forma de generación de la secuencia de randomización, el adecuado ocultamiento de la asignación de los pacientes a uno u otro grupo, el doble ciego y la descripción de los casos que abandonan el estudio, son elementos claves en la valoración metodológica de un RCT. La mala calidad metodológica de un RCT es un indicador de la posible presencia de sesgos en el proceso de investigación.

Linde *et al* describieron 4 predictores de buena calidad metodológica de los RCT homeopáticos: estar indexado en Medline, fecha de publicación 1990 o posterior, estar escrito en idioma inglés y que el tamaño de la muestra sea de 100 o mayor [27].

El tamaño también importa. El problema del azar

Intuitivamente sabemos que si arrojamus una moneda al aire tenemos un 50% de probabilidades de que salga cara. Si arrojamus una moneda al aire 4 veces, la probabilidad de obtener 4 caras es del 6,25%. Si la arrojamus 8 veces, la probabilidad de obtener 8 caras se reduce al 0,4%. Esperamos sacar 2 caras si arrojamus la moneda 4 veces y 4 si la arrojamus 8 veces. Pero obtener 4 caras arrojando la moneda 4 veces es todavía *probable*. Obtener 8 caras arrojando la moneda 8 veces es *muy improbable*. Las pruebas de significación estadística permiten inferir la probabilidad de que la diferencia observada entre muestras se deba meramente al azar. Si la probabilidad de que la diferencia observada se deba al azar es muy baja (en los RCT se acepta que *muy baja* es una probabilidad inferior al 5%) se asume que ésta se debe a otro u otros factores, como por ejemplo, a la intervención terapéutica. En el caso de la moneda, si la arrojamus 4 veces y obtenemos 4 caras, sabemos que la probabilidad de que esto ocurra por azar es del 6,25%, y podríamos afirmar que no se trata de un resultado estadísticamente significativo. En el segundo caso, 8 lanzamientos y 8 caras, la probabilidad de que esto ocurra por azar es tan baja (0,4%) que nos permitiría inferir que algo ocurre con la moneda, la cual podría estar trucada.

Un concepto importante referente a las diferencias que se observan por puro azar, es que éste puede alterar el resultado *en ambos sentidos*, contrariamente al sesgo que lo hace uno solo. Al arrojar la moneda 4 veces, existe un 6,25% de probabilidades de obtener 4 caras, pero también existe la misma probabilidad de obtener 4 cecas.

Las pruebas de significación estadística, algunas de las cuales son muy conocidas como el test t de Student, Chi cuadrado, test exacto de Fisher, la probabilidad binomial, etc., al calcular la probabilidad de que la diferencia encontrada entre dos grupos (p. ej., grupo tratado y grupo placebo) se deba al puro azar, toman en cuenta de manera determinante el tamaño de la muestra. Cuanto más grande es el tamaño de la muestra estudiada, menor la probabilidad de que por puro azar se detecte una diferencia significativa entre grupos cuando en realidad no la hay, o por el contrario, no se detecte una diferencia cuando en realidad sí la hay.

Si bien el tamaño de la muestra no constituye un sesgo por sí sólo, podría conducir indirectamente a un sesgo de publicación. De acuerdo a este último, tienen mayor probabilidad de ser publicados los RCT con un efecto terapéutico considerable y los RCT grandes aunque den un resultado negativo [28]. Como vimos, los RCT pequeños sólo por azar, tienen más chances de mostrar efectos terapéuticos muy positivos, aunque también tienen más chances de mostrar resultados muy negativos. O dicho de otra manera, suele existir una relación inversa entre el tamaño de los RCT y el efecto terapéutico medido.

Aunque por otro lado el médico homeópata e investigador Lex Rutten, además de sostener que este criterio de la importancia del tamaño no es universalmente compartido, explica que un RCT pequeño con pacientes mejor seleccionados puede dar un resultado más positivo sin que esto se deba necesariamente al azar. Esto podría

aplicarse a los RCT homeopáticos debido a que en el tratamiento homeopático clásico, en el proceso llamado de “individualización del caso” los medicamentos son seleccionados de acuerdo a las características de los pacientes que permiten pronosticar una mejor respuesta terapéutica [29].

El trabajo de Shang et al

Seleccionando arbitrariamente los trabajos

La comparación primaria de Shang *et al* es entre 110 RCT homeopáticos y 110 RCT convencionales, emparejados por enfermedad y medidas principales evaluadas (*main outcomes*). Los RCT homeopáticos seleccionados son el producto de una extensa actualización de la búsqueda hecha para un metanálisis previo [22]. Fue llevada a cabo en 19 bases de datos y a la selección anterior se agregaron RCT publicados entre 1995 y enero de 2003. De 170 RCT potencialmente elegibles, 60 fueron excluidos por diversos criterios preestablecidos.

Los RCT convencionales fueron extraídos del *Cochrane Controlled Trials Register*, el cual a principios de 2003 incluía más de 350.000 referencias bibliográficas. De acuerdo a Shang *et al*, “utilizaron números generados al azar por computadora para seleccionar uno entre varios RCT convencionales elegibles” emparejados para los mismos trastornos y las mismas medidas evaluadas (*outcomes*). Las medidas evaluadas en cada trabajo fueron seleccionadas de acuerdo a un criterio preestablecido. De acuerdo a los autores “las medidas evaluadas escogidas para cada trabajo y los trabajos fueron emparejados sin conocimiento de los resultados de los RCT”.

Algunas dudas metodológicas surgen de este procedimiento de selección de los trabajos [9]:

- ✓ ¿Son comparables 110 RCT homeopáticos seleccionados de entre 170, con 110 RCT convencionales extraídos de entre 350.000? La desproporción es evidente.
- ✓ Las referencias bibliográficas del *Cochrane Controlled Trials Register*, de donde fueron extraídos los RCT convencionales, provienen principalmente de las bases de datos de Medline y EMBASE (<http://www.thecochranelibrary.com/view/0/AboutTheCochraneLibrary.html#CENTRAL>). Ambas bases de datos indexan exclusivamente publicaciones revisadas por pares, que son las que en general aseguran una mayor calidad metodológica de los trabajos. ¿Son comparables entonces los RCT convencionales extraídos de estas bases de datos con los RCT homeopáticos extraídos de 19 bases de datos, si los niveles de exigencia para indexar una publicación difieren?
- ✓ Si bien Shang *et al* refieren haber seleccionado al azar uno de entre varios RCT convencionales elegibles para ser emparejados con sus pares homeopáticos, no describen cómo hicieron para seleccionar los “varios RCT elegibles” de los miles posibles.

Comparando peras con manzanas I

Las dudas planteadas arriba se vuelven aun más razonables cuando se comparan las características de los dos grupos de RCT seleccionados aportadas por Shang et al. Diferencias notorias surgen entre ambos (Tabla 1).

Tabla 1. Características de los 110 RCT homeopáticos y 110 RCT convencionales seleccionados por Shang et al

	110 RCT homeopáticos	110 RCT convencionales	P (2 colas) χ^2 *
Rango de año de publicación	1966-2003	1974-2002	
Idioma inglés (%)	53%	85%	< 0,0001
Publicación hecha en revistas (%)	85%	100%	< 0,0001
Publicación hecha en revista indexada en Medline (%)	41%	86%	< 0,0001

* Cálculos hechos con <http://www.vassarstats.net/>

De acuerdo a estos resultados, es de esperar una mejor calidad metodológica de los RCT convencionales y la comparación entre ambos grupos se vuelve muy cuestionable.

Pero los problemas de selección son aun mayores. Un análisis posterior de los datos aportados tardíamente por Shang et al [24] pudo constatar que:

- ✓ 16 (15%) de los RCT homeopáticos incluidos en el trabajo no se encontraban publicados contra 0% de los RCT convencionales, todos publicados. Un RCT no publicado tiene mayores chances de tener mala calidad metodológica y resultado negativo. Efectivamente, ninguno de estos 16 RCT tenía buena calidad metodológica y todos tenían resultado negativo.
- ✓ Fue excluido el RCT de Wiesenauer sobre poliartritis crónica [30], calificado como de buena calidad metodológica en el metanálisis de Linde [22] y "grande" (N= 176).
- ✓ Fueron excluidos otros 3 RCT homeopáticos con resultado positivo y buena calidad de acuerdo a análisis previos [31-33]. De acuerdo a las explicaciones brindadas por Shang et al 4 meses después de su publicación, estos 4 últimos RCT fueron excluidos porque no encontraron ningún RCT alopático con los cuales emparejarlos. Como se verá más adelante, el mencionado emparejamiento fue abandonado por Shang et al al sacar sus conclusiones finales, por lo que la exclusión de estos 4 RCT resulta como mínimo una ironía [4].
- ✓ Cuatro RCT homeopáticos, 3 de ellos con N >100, no fueron calificados como de buena calidad metodológica a pesar de haber sido considerados entre los 6 mejores en el metanálisis de Linde [34-37]. Esto podría mostrar que la calificación de la calidad metodológica no está exenta de cierta subjetividad [29].

Todos los elementos anteriores constituyen evidencia de *sesgo de selección*, y combinados influyen para que en conjunto los RCT homeopáticos muestren una peor calidad metodológica y un menor resultado terapéutico del que en realidad tienen.

Otras curiosidades en la selección de los RCT

Tres de los RCT convencionales seleccionados por Shang et al son acerca de tratamientos posteriormente prohibidos por la *Food and Drug Administration* (FDA) debido a su alta incidencia de efectos adversos riesgosos: dexfenfluramina en la obesidad, deladumone (andrógenos-estrógenos) en el amamantamiento y piroxicam para las lesiones de tejidos causadas por el amamantamiento. En el primer caso, la dexfenfluramina resultó altamente eficaz para bajar de peso mientras el tratamiento homeopático en el RCT correspondiente resultó ineficaz. ¿Cuál podría ser la utilidad práctica de comparar un tratamiento riesgoso con uno inefectivo? [24]

Sorpresa: los resultados no son los esperados

En contra de las predicciones y de los sesgos en la selección de los trabajos mencionados arriba, la primera y no menor de las sorpresas, es que los RCT homeopáticos resultan ser de mejor calidad metodológica que los convencionales. De acuerdo a Shang et al, 21 (19%) de los primeros contra 8 (9%) de los segundos fueron calificados como de la mejor calidad ($p=0,03$). Este resultado hace muy cuestionable cualquier comparación ulterior: es de esperar que al tener mejor calidad, los RCT homeopáticos arrojen un resultado terapéutico inferior a los convencionales.

Sin embargo y aun así, la mayoría de los RCT homeopáticos mostró un efecto terapéutico positivo y superior al placebo [Figura 1].

Figura 1. Diagrama de embudo de los 110 RCT homeopáticos (arriba) y 110 RCT convencionales (abajo) incluidos en el trabajos de Shang et al. *Odds ratio* <1 significa efecto terapéutico positivo. De Shang et al, con autorización.

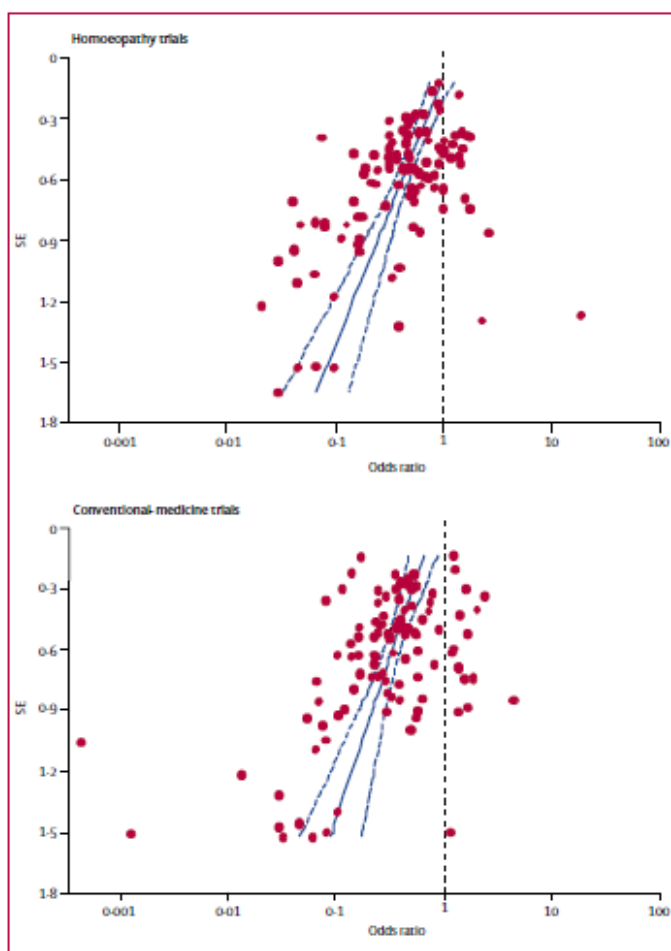


Figure 2: Funnel plot of 110 homeopathy trials and 110 matched conventional-medicine trials. Solid lines indicate predicted treatment effects from meta-regression, with dotted lines representing the 95% CI.

Esta figura, llamada diagrama de embudo (*funnel plot*), se debe interpretar de la siguiente manera: arriba se encuentran los resultados de los RCT homeopáticos y abajo los convencionales. Cada punto de color rojo representa el resultado terapéutico de un trabajo. En el eje de las X se encuentra la razón entre momios (*odds ratio*, OR ver *Glosario*) y en el eje de las Y el error estándar, una medida indirecta de los tamaños de las muestras. A menor error estándar, mayor el tamaño de las muestras. La línea vertical de puntos representa $OR=1$, ausencia de diferencia entre tratamiento y placebo. Cuanto menor es la OR (todos los puntos a la izquierda de la línea punteada) mayor es el efecto terapéutico.

Las figuras ponen en evidencia tres cosas: 1) La relación inversa entre el tamaño de la muestra y el efecto terapéutico detectado; 2) El hecho de que no adoptan la forma de un embudo invertido *simétrico*, hace sospechar tanto para los estudios convencionales como para los homeopáticos, algún grado de sesgo de publicación. Es decir, un número no determinable de RCT no publicados con resultados terapéuticos inferiores o nulos; 3) El gráfico de los RCT convencionales es más asimétrico que el de los RCT homeopáticos, evidenciando un mayor sesgo de publicación en los primeros.

Frente a este resultado, Shang et al, curiosamente se limitan a afirmar que “la mayoría de los OR indican un efecto beneficioso de la intervención”, sin hacer la cuantificación correspondiente. A simple vista y de manera global, no parece surgir una diferencia entre la efectividad del tratamiento homeopático y el convencional.

Gracias a la reconstrucción llevada a cabo por Rutten (ver abajo), una hipótesis fundamental de Shang et al pudo ser falsificada: que “los trabajos más grandes tienen mayor probabilidad de tener mejor calidad metodológica y menor sesgo que los chicos. Los efectos clínicos positivos observados a favor de la homeopatía en metanálisis previos podrían deberse a una exagerada influencia en los resultados de los RCT pequeños y por la tanto de menor calidad”. El 16,9% de los RCT homeopáticos y el 2,6% de los RCT convencionales con $N < 100$ resultaron ser de la mejor calidad metodológica. Conclusión: los RCT homeopáticos pequeños también resultaron tener mejor calidad metodológica ($P= 0,003$) [24, 38].

En resumen, considerando la totalidad de los trabajos seleccionados y a pesar de los sesgos de selección señalados más arriba, el tratamiento homeopático muestra ser como mínimo tan efectivo como el convencional para la mayoría de las enfermedades estudiadas. Aunque, como bien señalan Frass et al, no es correcto hablar de “comparación” entre tratamiento homeopático y convencional, ya que realidad y en contradicción con el título del trabajo de Shang et al, se trata de una investigación independiente de ambos métodos [15].

Además una pregunta incómoda e inquietante queda en el aire: ¿la investigación clínica es superior en calidad metodológica en el ámbito de la homeopatía que en el de la medicina convencional?

El problema de la heterogeneidad

La heterogeneidad en un metanálisis se define como la variabilidad estadística entre los RCT que se encuentran incluidos. Como es de esperar y debido a que ambos metanálisis comparan RCT de enfermedades e intervenciones variadas, Shang et al

observan una alta heterogeneidad tanto entre los RCT convencionales como entre los RCT homeopáticos.

Por si misma una heterogeneidad elevada cuestiona la validez de la combinación de resultados que se lleva a cabo en un metanálisis. La heterogeneidad afecta negativamente el resultado de un metanálisis de la siguiente manera: en un metanálisis, y especialmente en el modelo llamado de *efectos aleatorios* que es el correctamente utilizado por Shang et al en este caso, los cálculos matemáticos permiten establecer un *efecto combinado* y un *intervalo de confianza (IC) combinado* (para dicho efecto) de los RCT intervinientes. Pero cuánto más heterogéneos son los RCT entre si y menor su número, más ancho será el *IC combinado* resultante, disminuyendo las probabilidades de encontrar una diferencia estadísticamente significativa aunque esta en realidad exista. Dicho de otra manera, aumenta el riesgo de encontrarse con un falso negativo.

Pero hay algo aun más llamativo en los resultados de Shang et al: los RCT convencionales resultaron ser significativamente más heterogéneos ($I^2= 77\%$, heterogeneidad elevada) que los RCT homeopáticos ($I^2= 65\%$, heterogeneidad moderada), con una probabilidad de que la diferencia se deba al azar $p= 0,011$. Se puede ver esto a simple vista en la Figura 1: los RCT homeopáticos están mejor agrupados en el gráfico, mientras que los convencionales están más dispersos. Esta diferencia en la heterogeneidad es otro elemento que hace cuestionable una comparación de ambos grupos de RCT.

Comparando peras con manzanas II

Llegado a este punto, el trabajo de Shang et al ya podría darse por concluido. Por diversos motivos los RCT homeopáticos y convencionales ya no pueden compararse: hay sesgos de selección evidentes y los RCT homeopáticos, tanto grandes como pequeños, son de mejor calidad metodológica, menos heterogéneos entre si y parecerían sufrir un menor sesgo de publicación. La mayoría de los RCT homeopáticos y convencionales muestran un efecto positivo.

Pero a continuación y sorprendentemente, Shang et al deciden limitar el análisis a los RCT “grandes y de buena calidad metodológica” y *sub-seleccionan* 8 RCT homeopáticos y 6 RCT convencionales.

Inmediatamente surgen varios problemas serios que los autores no parecen reconocer:

- ✓ Ya no se están emparejando los RCT para la misma afección y mismas medidas evaluadas como estaba propuesto inicialmente. Sólo quedan 2 RCT (uno homeopático y uno convencional) de tratamiento de la diarrea infantil y 3 RCT (1 homeopático y 2 convencionales) de tratamiento de la influenza que pueden ser emparejados. Seis RCT homeopáticos y 3 RCT convencionales quedan sin emparejar de acuerdo a la enfermedad estudiada. Esto es un abandono completo de la propuesta inicial y una contradicción con el mismo título del trabajo, lo que lleva a dudar de que haya una estrategia planteada *a priori* o un análisis hecho *a posteriori* ante el resultado inicial negativo y en desacuerdo con la expectativas de los investigadores. (Ver más abajo *análisis post hoc*) [24].

- ✓ Al ser tan pocos los trabajos escogidos (8 homeopáticos y 6 convencionales), y casi todos de enfermedades diferentes, la heterogeneidad aumenta haciendo aún más cuestionable su combinación en un metanálisis y aumentando la probabilidad de obtener un resultado falso negativo.
- ✓ ¿Cuán grande es más grande? Shang et al se limitan a informar que seleccionaron los RCT incluidos en el cuartil inferior del error estándar sin esgrimir las razones. La reconstrucción de Rutten permitió detectar que este valor de corte para la definición de “más grande” introduce una nueva desigualdad entre los RCT homeopáticos y convencionales. El N para los primeros es de 98 para los primeros y de 146 para los segundos.

Es en el análisis de estos pequeños dos subgrupos que se basa la conclusión de Shang et al. Los 8 RCT homeopáticos dan como resultado una OR= 0,88 (IC 95%: 0,65–1,19), positivo pero no significativo desde el punto de vista estadístico, mientras que los 6 RCT convencionales arrojan una OR= 0,58 (IC 95%: 0,39–0,85).

Es sorprendente que Shang et al “olvidaran” referenciar estos pocos RCT seleccionados, de manera que no se pudo saber en qué estudios se basaba este supuesto “fin de la homeopatía” hasta unos 4 meses más tarde. Recién publicaron estas referencias después de recibir las críticas al respecto.

¿Cuáles son estos RCT que “demuestran” que la homeopatía es placebo?

En un metanálisis, el cálculo matemático de los efectos combinados no exige de la revisión libre y crítica de cada uno de los estudios individuales [23]. Es ilustrativo examinar cuáles son estos 8 RCT en que se basa la conclusión de Shang et al (Tabla 2).

Tabla 2. Ocho RCTs homeopáticos seleccionados por Shang et al para arribar a las conclusiones finales de que “la homeopatía es placebo”

Autor	Objeto	Medicamento	N	Resultado	Comentarios
Rottey EED <i>et al</i> , 1995 [39]	Prevención de síntomas de la influenza	Complejo homeopático conteniendo virus de influenza y varias bacterias en dilución 200K o placebo	50 1	Positivo y estadísticamente significativo de acuerdo a una escala de Likert. Número de síntomas gripales con una diferencia a favor del grupo tratado pero no significativa estadísticamente	Probablemente Shang <i>et al</i> hayan tomado la segunda variable para el análisis
Vickers AJ <i>et al</i> , 1998 [40]	Tratamiento del dolor muscular en maratonistas	<i>Arnica montana</i> 30d o placebo	40 0	Inefectivo	Clínicamente irrelevante. Otros RCT publicados sobre el mismo tema también dan resultado negativo. Este trabajo influye decisivamente en el resultado combinado del

					metanálisis.
Papp R <i>et al</i> , 1998 [41]	Tratamiento de cuadros gripales	Oscillococcinum 200K o placebo	37 2	Positivo y estadísticamente significativo	
Labrecque M <i>et al</i> , 1992 [42]	Tratamiento de las verrugas plantares	Complejo con <i>Thuja occidentalis</i> 30C, <i>Antimonium crudum</i> 5c y <i>Nitricum acidum</i> 7c o placebo	16 2	Sin diferencia entre tratamiento y placebo	Los complejos no forman parte de la práctica habitual de los homeópatas
Jacobs J <i>et al</i> , 2000 [43]	Tratamiento de la diarrea infantil aguda	Medicamento individualizado en dilución 30c o placebo	12 6	Positivo y estadísticamente significativo	
Weiser M <i>et al</i> , 1994 [44]	Tratamiento de la sinusitis crónica	Complejo conteniendo <i>Euphorbium</i> 6d, <i>Pulsatilla pratensis</i> 3d, <i>Luffa operculata</i> 3d, <i>Hepar sulphur</i> 10d o placebo	10 2	Sin diferencia entre tratamiento y placebo	Los complejos no forman parte de la práctica habitual de los homeópatas
Walach H <i>et al</i> , 1997 [45]	Tratamiento de la cefalea	Medicamento individualizado en diluciones c o LM, o placebo	98	Sin diferencia entre tratamiento y placebo	
Schmidt JL, Ostermayr B, 2002 [46]	Efecto sobre el índice de masa corporal en pacientes que hacen ayuno	<i>Thyroidinum</i> 30C o placebo	20 8	Sin diferencia entre tratamiento y placebo	No es forma habitual de prescripción homeopática.

En resumen:

- ✓ Sólo dos estudios que representan la homeopatía tradicional, es decir, la que utiliza un solo medicamento homeopático bien seleccionado para el caso: uno con resultado positivo (Jacobs, diarrea infantil) y uno con resultado negativo (Walach, cefalea).
- ✓ Tres estudios con complejos, uno positivo (Rottey, mezcla de nosodes para la prevención de la gripe) y dos negativos (Labrecque, verrugas plantares y Weiser, sinusitis).
- ✓ Tres estudios de los llamados “clínicos”: a todos los pacientes el mismo medicamento simple. Uno positivo (Papp, Oscillococcinum, cuadros gripales) y dos negativos (Vickers, *Arnica montana* en dolorimiento muscular post-ejercicio y Schmidt, *Thyroidinum* en el índice de masa corporal).

Es muy difícil justificar la idea de que esta selección resulte representativa de la práctica habitual de la homeopatía, porque en realidad no lo es [11].

Pero hay algo más. Supongamos que como sostienen Shang *et al* efectivamente han encontrado 8 RCT homeopáticos de muy buena calidad metodológica y grandes, de manera que la probabilidad de que en ellos haya sesgo o diferencias debidas al azar sea muy baja. Aun así, 3 de los 8 dan un resultado positivo y estadísticamente

significativo. Esto es, menos de 5% de probabilidades de que las diferencias encontradas se deban al azar. Para hacernos una idea de lo que esto significa, imaginemos tener 8 bolsas, dentro de cada una de las cuales hay 95 bolillas blancas y 5 rojas. A ciegas sacamos una bolilla de cada una de ellas. ¿Cuál sería la probabilidad de sacar 3 bolillas rojas? Bajísima: menor de 1 en 1000 (cálculo de probabilidad binomial, <http://www.vassarstats.net/>). Por supuesto, esta forma de hacer estadística no es correcta. Pero la pregunta queda en pie: ¿cómo es posible que de 8 RCT homeopáticos de muy buena calidad metodológica y N grande, 3 den resultado positivo y estadísticamente significativo? Si el tratamiento homeopático fuera en efecto similar al placebo, ¿no sería este hallazgo tan improbable como casi imposible? Y al fin y al cabo, que un porcentaje de los RCT homeopáticos den un resultado negativo, ¿es prueba suficiente de que la homeopatía como un todo es inefectiva? Además ¿dónde quedó la sentencia que afirma que la falta de prueba a favor no es igual a prueba en contrario?

El problema de Arnica en el dolorimiento muscular

El caso de *Arnica* presenta ribetes de especial interés. El dolorimiento muscular post-ejercicio ha sido una de las indicaciones populares de *Arnica* utilizado homeopáticamente. Varios RCT más pequeños evaluando este posible efecto de *Arnica* habían sido publicados en el pasado, con resultados no concluyentes. El estudio de Vickers es probablemente el RCT definitivo al respecto: *Arnica* es inefectivo para el dolorimiento muscular post-ejercicio [40].

El problema reside en que al ser este RCT un estudio grande y de buena calidad y con resultado negativo, arrastra negativamente el resultado combinado de Shang *et al.* La pregunta es, ¿tiene sentido incluir en el análisis un estudio RCT negativo, que es la confirmación de un resultado que ya se sospechaba y que además se refiere a un problema clínicamente irrelevante?

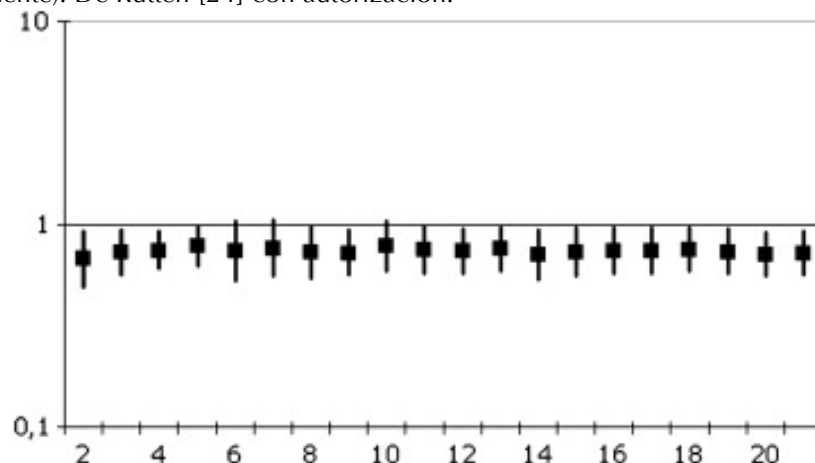
Análisis de sensibilidad: un ejercicio necesario e interesante con resultados sugestivos

En un análisis de sensibilidad, elemento fundamental de todo metanálisis y omitido por Shang *et al.*, se evalúa qué ocurre con los resultados del metanálisis si se excluyen uno a uno los estudios seleccionados. Esto hicieron de diversas maneras Rutten y Lüdtke [24, 38] y obtuvieron resultados por demás elocuentes:

- ✓ Haciendo el análisis combinado de los 21 RCT homeopáticos señalados por Shang *et al.* como de la mejor calidad, el resultado no es compatible con la idea de que no hay efectos homeopáticos específicos sino sólo efecto placebo: OR= 0,76 (IC 95%: 0,59 – 0,99; p= 0,039).
- ✓ Si se consideran los 21 RCT homeopáticos de mejor calidad metodológica detectados por Shang *et al.* y se los coloca según número decreciente de N, y a continuación se selecciona cualquier número de trabajos entre 14 y 21 (con la sola excepción de 17) el resultado combinado también es a favor del tratamiento homeopático y estadísticamente significativo.

- ✓ Cuatro de estos 21 RCT son estudios de *Arnica* en dolorimiento muscular post-ejercicio. Un metanálisis de estos 4 RCT muestra que *Arnica* no funciona para esta indicación. Si se quitan estos 4 estudios del análisis, los 17 estudios restantes muestran un efecto combinado muy marcado a favor de la homeopatía (OR=0,68; IC 95%: 0,52 – 0,90; p=0,007).
- ✓ El efecto positivo se mantiene si ahora se seleccionan sólo los 8 RCT más grandes. Es más, la heterogeneidad se reduce considerablemente en las dos situaciones anteriores ($I^2= 52,1\%$ y $37,1\%$ respectivamente) sugiriendo que los trabajos de *Arnica* contribuyen con gran parte de la misma así como al resultado combinado negativo.
- ✓ Si se eliminan los RCT sobre *Arnica* en el dolorimiento muscular post-ejercicio y se agregan los 4 RCT considerados por Linde como de la mejor calidad y excluidos por Shang et al y se aumenta uno a uno el número de estudios incluidos en orden de N decreciente comenzando por los 2 más grandes, se observa que cualquiera sea el número de RCT incluidos entre 2 y 21, con la sola excepción de 6, 7 o 10, el resultado combinado es favorable a la homeopatía y estadísticamente significativo (Figura 2).
- ✓ Si se hubiera elegido la mediana de N como valor de corte (N= 65, en ambos grupos) y excluyendo el RCT de dolorimiento muscular post-ejercicio (ver más abajo), los RCT homeopáticos de la mejor calidad habrían quedado incluidos en el análisis y el resultado combinado habría sido altamente positivo para la homeopatía: OR= 0,66 (95% IC 0,49–0,89).

Figura 2. OR (IC 95%) combinada acumulada cuando se incrementa uno a uno el número de RCT de la más alta calidad de acuerdo a N decreciente, comenzando con los dos más grandes. A los 21 RCT homeopáticos de buena calidad detectados por Shang et al se agregaron los RCT calificados de buena calidad por Linde y excluidos por Shang et al, y se excluyeron los RCT de *Arnica* en dolorimiento muscular post-ejercicio. Todas las opciones son favorables a la homeopatía con excepción de tres (6, 7 y 10 RCT, con valores de corte de N= 144, 122 y 98 respectivamente). De Rutten [24] con autorización.



Rutten concluye que la sub-selección de estos 8 RCT homeopáticos es, de todas las sub-selecciones posibles, la que da el resultado más desfavorable a la homeopatía [30]. También se pregunta qué pasaría si la homeopatía funcionara en

algunas enfermedades y no en otras. ¿Qué sentido tendría hacer un metanálisis de unos pocos RCT en este contexto? [24]

Los autores de estos análisis de sensibilidad afirman con objetividad que debido a que este estudio es *post hoc*, no permite sacar conclusiones a favor de la homeopatía. Pero podría considerarse una prueba a favor de que la selección de Shang et al de los 8 RCT homeopáticos, la cual es muy notablemente conveniente para arribar a las conclusiones de que la homeopatía es placebo, también haya sido hecha *post hoc* y que por lo tanto fuera inválida.

La insólita conclusión extraída de 8 RCT homeopáticos sobre enfermedades respiratorias, de buena calidad y con resultado positivo

En su búsqueda de RCT homeopáticos, Shang et al encontraron 8 RCT sobre tratamiento de enfermedades de vías aéreas superiores que no presentan evidencias de sesgo y que tienen un resultado combinado fuertemente positivo a favor de la homeopatía (OR= 0,36; IC 95%: 0,26-0,50). Este hallazgo debería conducir a la conclusión de que la homeopatía es efectiva en el tratamiento las enfermedades de vías aéreas superiores. Pero en su lugar, Shang et al afirman que “debido a los sesgos prevalentes en este tipo de publicaciones, lleva a pensar que los resultados de estos estudios no son confiables”.

Shang et al además sostienen que en un metanálisis de sólo 8 RCT es difícil detectar la presencia de sesgos y que este es otro motivo para desconfiar de los resultados de este metanálisis sobre homeopatía y enfermedades de las vías aéreas superiores. Pero, la conclusión de Shang et al de que la homeopatía es placebo, ¿no se basa precisamente en un metanálisis de 8 RCT homeopáticos? Entonces ¿son 8 RCT a veces suficientes y otras no? [13]

¿Análisis post hoc?

Un postulado básico de la investigación científica afirma que las hipótesis de investigación deben preceder a los datos. Si se procede de manera inversa, es decir, se elaboran las hipótesis a partir de los datos (*análisis post hoc*), se vuelve probable que se encuentre la manera de reordenar o de analizar estos últimos de forma que permitan “encontrar” los resultados deseados.

Diversos elementos de juicio hacen sospechar que Shang et al cambiaron las hipótesis iniciales al encontrarse con los datos. Algunos de estos son:

- ✓ Si planeaban de entrada comparar sólo RCT de buena calidad y grandes, ¿por qué no predefinieron los criterios de inclusión de estas categorías de manera de seleccionar exclusivamente RCT que cumplieran con estos requisitos? ¿Por qué seleccionar 110 RCT homeopáticos y 110 convencionales y hacer una comparación entre ellos, para después prácticamente ignorar los resultados y recurrir a una sub-selección?
- ✓ Si el título del trabajo dice explícitamente “*Comparative study of placebo-controlled trials of homoeopathy and allopathy*” (Estudio comparativo de RCT homeopáticos y convencionales), ¿cómo es posible que al final se abandone esta premisa presentando dos grupos de RCT que son incomparables, ya que son para enfermedades diferentes?

- ✓ ¿Cómo es posible que escogieran un valor de corte de N diferente para los RCT homeopáticos y los convencionales? Un valor de corte por cierto tan conveniente para llegar a los resultados deseados que cuesta creer que haya sido establecido *a priori*.
- ✓ ¿Cómo es posible que hayan omitido la referencia de los RCT seleccionados para justificar sus conclusiones?
- ✓ ¿Por qué omitieron hacer el análisis de sensibilidad?

¿Importa el tipo de homeopatía empleada?

Shang et al clasifican la estrategia homeopática empleada como sigue:

- ✓ *Homeopatía clínica*, 44% de los RCT, cuando no se realiza una historia clínica al paciente y todos reciben un mismo medicamento simple.
- ✓ *Complejos*, 32% de los RCT, que se refiere a la prescripción de mezclas de diversos medicamentos.
- ✓ *Homeopatía clásica*, 16% de los RCT, cuando a cada paciente se realiza una historia clínica comprehensiva y se le prescribe un único medicamento simple.
- ✓ *Isopatía*, 7% de los RCT, cuando se prescribe el mismo elemento que se considera la causa del trastorno.

Un análisis comparativo del efecto terapéutico hecho por Shang et al de estos 4 subgrupos no mostró diferencias significativas entre ellos. Pero como no fueron descritas las características de estos 4 sub-grupos de RCT, probablemente carezca de significado.

Esta desproporción encontrada en contra de la *homeopatía clásica*, a pesar de que probablemente sea la estrategia preferida por los homeópatas, se explicaría por varias causas. La primera, que el diseño de un RCT de homeopatía clásica es mucho más complejo y costoso que para las otras 3 formas de homeopatía. La segunda, que la investigación de los complejos se encuentra muchas veces impulsada por el interés comercial de una compañía farmacéutica mientras que los llamados medicamentos homeopáticos unitarios carecen de incentivos económicos por tratarse de genéricos sin patente.

Ocultando información y la obligada reconstrucción del análisis de Shang et al

Shang et al fueron inmediatamente criticados porque en su trabajo ocultaron información clave, impidiendo hacer una evaluación minuciosa. Los datos faltantes fueron revelados 4 meses más tarde y estos eran: 1) Las referencias de los de RCT homeopáticos excluidos; 2) Las referencias de los RCT homeopáticos considerados como de mejor calidad metodológica; 3) Las referencias de los 8 RCT homeopáticos y los 6 RCT convencionales que llevaron a la conclusión final [24].

Otros datos esenciales y necesarios para la valoración de un metanálisis fueron directamente negados a los investigadores que los solicitaron directamente [30]. La falta de transparencia fue uno de los cuestionamientos éticos más repetidos entre los críticos de Shang et al.

El trasfondo

En 1999, varias formas de medicina complementaria incluyendo la homeopatía fueron provisionalmente incorporadas al sistema de seguridad social suizo. Mientras esto ocurría, se llevaba a cabo una exhaustiva investigación para evaluar la efectividad y la relación costo-beneficio de estas formas de medicina. El programa fue conocido como *Programm Evaluation Komplementärmedizin* (PEK) y una de las instituciones involucradas era el *Institut für Sozial und Preventiv Medizin* (ISPM). Este último es el que llevó a cabo el estudio criticado en esta publicación y 7 de los 8 autores del trabajo pertenecían a él en el momento de la publicación. Las conclusiones ya habían sido anunciados a la prensa en agosto de 2003, 2 años antes de la publicación del trabajo, en una acción éticamente dudosa: al no estar publicado el trabajo resultaba imposible conocer cómo la investigación había sido hecha. Los métodos del trabajo fueron muy criticados por la misma dirección del PEK, destacando una obviedad: enfermedades y tratamientos diferentes no pueden ser abordados estadísticamente como si fueran un solo tratamiento para una sola enfermedad.

El consejo internacional de revisión del PEK se quejó públicamente de “interferencias políticas en el proceso científico” [13]. Algunos elementos no científicos se mezclan en la discusión científica. Matthias Egger, cabeza del ISPM y uno de los autores del trabajo, es conocido desde hace años por una firme postura contraria a la homeopatía. ¿Puede esta predisposición haber influido en el desarrollo del trabajo en cuestión? [4, 11]

Rutten se pregunta por qué una revista médica como *The Lancet* pudo publicar un trabajo de investigación que no reúne los criterios mínimos exigidos por los editores. Especula con que en la época circulaba discretamente para ser evaluado el borrador de un reporte de la Organización Mundial de la Salud (OMS) cuya conclusión final era que la homeopatía es efectiva y recomendaba más investigación en este campo. De hecho y curiosamente, en el mismo número de *The Lancet* en que se publica el trabajo de Shang et al, se publica también una fuerte crítica, hostil y casi intimidatoria sobre este borrador de la OMS, que evidentemente se había filtrado rompiendo con elementales reglas éticas de reserva [47].

El trabajo de Shang et al era sólo uno de los trabajos encomendados por el PEK. Otras investigaciones en marcha mostraban que los pacientes tratados con homeopatía tenían mejores respuestas terapéuticas, menos internaciones hospitalarias e incurrían en menores costos económicos. Todas estas líneas de investigación fueron suspendidas y se prohibió publicar los resultados [4].

La supuesta no plausibilidad de la homeopatía

La cuestión de la plausibilidad de la acción biológica de las altas diluciones utilizadas en homeopatía es un punto central en la discusión. Ya había sido señalada de manera muy elocuente por los autores de la primera revisión sistemática de RCT homeopáticos publicada en 1991 de la siguiente forma: “La cantidad de evidencia positiva aun entre los mejores trabajos fue una sorpresa para nosotros. Basados en esta

evidencia estaríamos listos para aceptar que la homeopatía puede ser eficaz si sólo el mecanismo de acción fuera más plausible” [48].

Pero como señala Fisher, la plausibilidad está en la mente del observador y puede cambiar de una generación a otra. Aunque no es objeto de este trabajo, es evidente que Shang et al fallan al evitar considerar la creciente y cada vez más refinada evidencia de la acción biológica de las altas diluciones que proviene de las ciencias básicas [18].

Conclusiones

Del trabajo de Shang et al y los análisis críticos hechos por varios autores, pueden desprenderse las siguientes conclusiones:

- ✓ La mayoría de los RCT homeopáticos analizados muestran un resultado clínicamente positivo y estadísticamente significativo.
- ✓ La homeopatía es eficaz en el tratamiento de las afecciones de vías aéreas superiores.
- ✓ *Arnica* es ineficaz en el tratamiento del dolor muscular post-ejercicio físico.
- ✓ La efectividad de la homeopatía se mantiene si el análisis se limita a los estudios de mejor calidad metodológica.
- ✓ Los RCT homeopáticos son de mejor calidad metodológica que sus pares convencionales.
- ✓ La mayoría de los RCT convencionales analizados muestran un resultado clínicamente positivo y estadísticamente significativo.

Los RCT homeopáticos y convencionales que analizan Shang et al son incomparables entre sí:

- ✓ La calidad metodológica es mayor en los RCT homeopáticos.
- ✓ La heterogeneidad es menor en los RCT homeopáticos.
- ✓ Los criterios para seleccionar uno y otro grupo no son similares.
- ✓ Tres RCT convencionales evalúan tratamientos no permitidos por sus efectos adversos, lo que carece de sentido.

Más incomparables aún son los 8 RCT homeopáticos y 6 RCT convencionales seleccionados al final:

- ✓ Hay elementos de juicio que permiten suponer que esta sub selección fue hecha *post hoc*.
- ✓ Los RCT homeopáticos y convencionales no se encuentran emparejados por enfermedad como había sido propuesto inicialmente.
- ✓ El N de ambos grupos es diferente.

La conclusión de Shang et al basada en 8 RCT homeopáticos de que los efectos clínicos de la homeopatía observada en los estudios son efectos placebo no es válida:

- ✓ Como se menciona arriba, esta sub-selección podría haber sido hecha *post hoc*.

- ✓ Se trata de 8 trabajos acerca de 8 enfermedades diferentes y la heterogeneidad es muy elevada.
- ✓ Uno de los RCT (*Arnica* en el dolor post ejercicio físico) es el estudio que de manera definitiva muestra la ineficacia del tratamiento con este medicamento de este problema.
- ✓ Estos 8 RCT no son representativos de la práctica habitual de los homeópatas.
- ✓ No se puede excluir un resultado falso negativo.

Varios sesgos y problemas metodológicos quedan en evidencia en la publicación de Shang et al:

- ✓ Sesgos en la selección de los RCT homeopáticos que combinados tienden a disminuir su calidad y efectividad.
- ✓ Posible análisis post hoc de los datos, para favorecer las conclusiones esperadas.
- ✓ Omisión del análisis de sensibilidad.

En definitiva, se puede afirmar que las conclusiones de Shang et al no solamente no son válidas sino que apuntan en dirección contraria: *los resultados clínicos de los RCT homeopáticos no son compatibles con la idea de que la homeopatía es placebo.*

Glosario

✓ **Ensayos clínicos randomizados y controlados** (RCT, de *randomized controlled trials*): En un RCT se comparan los resultados obtenidos en dos grupos de pacientes seguidos concomitantemente (en paralelo), uno que recibe el tratamiento investigado (activo o verum) y otro (*grupo control*) que recibe un placebo indistinguible del tratamiento verdadero (o un tratamiento diferente). Los pacientes estudiados, que siempre constituyen una muestra del universo de pacientes posibles escogidos de acuerdo a criterios preestablecidos que permitan minimizar el impacto de la presencia de elementos confundentes, son asignados a uno u otro grupo al azar (randomización o aleatorización). Ni ellos ni los médicos tratantes saben si reciben tratamiento verdadero o placebo (doble ciego). La existencia de un grupo control permite diferenciar con claridad el efecto del tratamiento del de otros factores que pueden influir en el curso de la enfermedad. La randomización asegura razonablemente que esos otros factores que podrían influir en el curso de la enfermedad se distribuyan uniformemente entre los dos grupos y de que todos los pacientes tengan las mismas chances de recibir tratamiento o placebo. El doble ciego permite que el efecto placebo se distribuya uniformemente entre los dos grupos y que la subjetividad de los pacientes y médicos no influya en la evaluación de los resultados.

✓ **Odds Ratio (OR) o razón entre momios**: Es una forma de cuantificar la diferencia observada en el tratamiento de dos grupos. OR= 1 significa que no hay diferencias entre grupos. De acuerdo a como fueron hechos los cálculos por Shang et al y para el resto de las referencias en esta publicación, OR <1 significa una diferencia

a favor del grupo tratado en comparación con el grupo placebo. Por ejemplo, OR= 0,5 implica que el tratamiento de 2 veces más efectivo que el placebo.

✓ **Intervalo de confianza de 95%** (IC95%): Es otra manera de expresar la significancia estadística. Expresa los valores entre los cuales, repitiendo la experiencia un número de veces en las mismas condiciones, se encontrará el resultado el 95% de las veces. Otra manera de decirlo es que existe un 95% de probabilidades de que el verdadero valor se encuentre comprendido por dicho intervalo. Para todas las referencias de este trabajo, si el IC 95% comprende el valor 1, la diferencia hallada no es estadísticamente significativa.

✓ **P:** Es la probabilidad de que la diferencia encontrada se deba meramente al azar. Se suele expresar en números absolutos. P= 0,04 significa P= 4%.

✓ **N:** es el número de sujetos sobre los cuales se lleva a cabo la experimentación.

Referencias bibliográficas

- 1- Shang A, Huwiler-Müntener K, Nartey L, Jüni P, Dörig S, Sterne JAC, et al. Are the clinical effects of homeopathy placebo effects? Comparative study of placebo-controlled trials of homeopathy and allopathy. *Lancet* 2005;366(9487):726–732.
- 2- The end of homeopathy. *Lancet* 2005;366(9487):690.
- 3- Linde K, Jonas W. Are the clinical effects of homeopathy placebo effects? *Lancet* 2005;366(9503):2081–2082.
- 4- Rutten L, Stolper E. "Proof" against homeopathy in fact supports Homeopathy Homeopathy. 2006;95(1):57–61.
- 5- Skandhan K, Amith S, Avni S. Are the clinical effects of homeopathy placebo effects? *Lancet* 2005;366(9503):2085.
- 6- Dantas F. Are the clinical effects of homeopathy placebo effects? *Lancet* 2005;366(9503):2083; author reply 2083–2086.
- 7- Fisher P, Berman B, Davidson J, Reilly D, Thompson T. Are the clinical effects of homeopathy placebo effects? *Lancet* 2005;366(9503):2082–2083.
- 8- Rosenbaum P. What end is this anyway? *Homeopathy* 2006;95(1):55–56.
- 9- Reilly D. Sir? Is that bias? *Homeopathy* 2006;95(1):55.
- 10- Raoult D. Are the clinical effects of homeopathy placebo effects? *Lancet* 2005;366(9503):2085–2086.
- 11- Thurneysen A. Open letter to the editor of The Lancet from the Swiss Association of Homeopathic Physicians (SVHA). *Homeopathy* 2006;95(1):61–62.
- 12- Walach H, Jonas W, Lewith G. Are the clinical effects of homeopathy placebo effects? *Lancet* 2005;366(9503):2081; author reply 2083–2086.
- 13- Fisher P. Homeopathy and The Lancet. *Evid Based Complement Alternat Med* 2006;3(1):145–147.
- 14- Poitevin B, Demonceaux A, Jeulin D, Popowski P, Sarembaud A, Guermonprez M, et al. French doctors' response. *Homeopathy* 2006;95(1):63–64.

- 15- Frass M, Schuster E, Muchitsch I, Duncan J, Gei W, Kozel G, et al. Bias in the trial and reporting of trials of homeopathy: a fundamental breakdown in peer review and standards? *J Altern Complement Med* 2005;11(5):780–782.
- 16- Albonico H, Bruno Ferroni B. Etude sur l'homéopathie: anatomie d'une opération statistique. *Bulletin des médecins suisses* 2006;45: 1954–1957.
- 17- Bell IR. All evidence is equal, but some evidence is more equal than others: can logic prevail over emotion in the homeopathy debate? *J Altern Complement Med* 2005;11(5):763–769.
- 18- Fisher P. Changing reality. *Homeopathy* 2006;95(1):1–2.
- 19- Peters D, Shang et al. carelessness, collusion, or conspiracy? *J Altern Complement Med* 2005;11(5):779–780.
- 20- Kiene H, Kienle GS, Von Schön-Angerer T. Failure to exclude false negative bias: a fundamental flaw in the trial of Shang et al. *J Altern Complement Med* 2005;11(5):783.
- 21- Fletcher RH, Fletcher SW. *Epidemiología clínica*, 4a ed. Barcelona: Wolters Kluwer Health España, Lippincott Williams & Wilkins; 2008.
- 22- Linde K, Clausius N, Ramirez G, Melchart D, Eitel F, Hedges LV, et al. Are the clinical effects of homeopathy placebo effects? A meta-analysis of placebo-controlled trials. *Lancet*.1997;350(9081):834–843.
- 23- Rutten ALB, Stolper CF. The 2005 meta-analysis of homeopathy: the importance of post-publication data. *Homeopathy* 2008;97(4):169–177.
- 24- Shang A, Jüni P, Sterne JA, Huwiler-Müntener K, Egger M. Are the clinical effects of homeopathy placebo effects? Authors' reply. *Lancet* 2005;366(9503):2083–2085.
- 25- Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273(5):408–412.
- 26- Linde K, Scholz M, Ramirez G, Clausius N, Melchart D, Jonas WB. Impact of study quality on outcome in placebo-controlled trials of homeopathy. *J Clin Epidemiol* 1999;52(7):631–636.
- 27- Linde K, Jonas WB, Melchart D, Willich S. The methodological quality of randomized controlled trials of homeopathy, herbal medicines and acupuncture. *Int J Epidemiol* 2001;30(3):526–531.
- 28- Baigent C. The need for large-scale randomized evidence. *Br J Clin Pharmacol* 1997 [cited January 27, 2013];43(4):349–353.
- 29- Rutten ALB. The 2005 Lancet review proved superior quality of homeopathy trials. Opening lecture at the LMHI Congress 2009. Warsaw. Retrieved from: <http://www.dokterrutten.nl/collega/Liga09.pdf> Último acceso: Maio 2013.
- 30- Wiesenauer M, Gaus W. Wirksamkeitsnachweis eines Homöopathikums bei chronischer Polyarthrit. Eine randomisierte Doppelblindstudie bei niedergelassenen Ärzten. *Aktuelle Rheumatologie* 1991;16:1–9.
- 31- Arnal-Laserre MN. Préparation à l'accouchement par homéopathie. Paris: Université René Descartes; 1986.
- 32- Maiwald L, Weinfurter T, Mau J, Connert WD. Therapie des grippalen Infektsy. *Arzneimittelforschung* 1988; 38: 578–582.
- 33- Fisher P, Greenwood A, Huskisson EC, et al. Effect of homeopathic treatment on fibrositis (primary fibromyalgia). *Br Med J* 1989; 299: 365–366.

- 34- de Lange de Klerk ESM. Effects of homeopathic medicines on children with recurrent upper respiratory tract infections. *BMJ* 1994; 309: 1329–1332.
- 35- Reilly DT, Taylor MA, McSharry C, Aitchison T. Is homeopathy a placebo response? Controlled trial of homeopathic potency with pollen in hay fever as model. *Lancet* 1986;881–886.
- 36- Hofmeyr GJ, Picconi V, Blauhof P. Postpartum homeopathic *Arnica montana*: a potency-finding pilot study. *Br J Clin Pract* 1990; 44:619–621.
- 37- Reilly D, Taylor MA, Beattie NG, et al. Is evidence for homeopathy reproducible? *Lancet* 1994; 344:1601–1606.
- 38- Lüdtke R, Rutten ALB. The conclusions on the effectiveness of homeopathy highly depend on the set of analyzed trials. *J Clin Epidemiol* 2008;61(12):1197–1204.
- 39- Rottey EED, Verleye GB, Liagre RLP. Het effect van een homeopathische bereiding van micro-organismen bij de preventie van griepsymptomen - Een gerandomiseerd dubbel-blind onderzoek in de huisartspraktijk. *Tijdschr Integ Geneeskunde* 1995; 11: 54–58.
- 40- Vickers AJ, Fisher P, Wyllie SE, Rees R. Homeopathic *Arnica* 30X is ineffective for muscle soreness after long-distance running – A randomized, double-blind, Placebo-controlled trial. *Clin J Pain* 1998; 14(3): 227–231.
- 41- Papp R, Schuback G, Beck E, et al. *Oscillococcinum* in patients with influenza-like syndroms: A placebo controlled double-blind evaluation. *Brit Homeopath J* 1998; 87(2): 69–76.
- 42- Labrecque M, Audet D, Latulippe L, Drouin J. Homeopathic treatment of plantar warts. *Can Med Assoc J* 1992; 146(10): 1749–1753.
- 43- Jacobs J, Jiménez LM, Malthouse S, et al. Homeopathic treatment of acute childhood diarrhea - Results from a clinical trial in Nepal. *J Alternat Complement Med* 2000; 6(2): 131–139.
- 44- Weiser M, Clasen BPE. Randomisierte plazebokontrollierte Doppelblindstudie zur Untersuchung der klinischen Wirksamkeit der homöopathischen *Euphorbium compositum*-Nasentropfen S bei chronischer Sinusitis. *Forsch Komplementarmed* 1994; 1: 251–259.
- 45- Walach H, Haeusler W, Lowes T, et al. Classical homeopathic treatment of chronic headaches. *Cephalalgia* 1997; 17: 119–126.
- 46- Schmidt JM, Ostermayr B. Does a homeopathic ultramolecular dilution of *Thyroidinum* 30CH affect the decrease of body weight reduction in fasting patients? A randomised Placebo-controlled double-blind trial. *Homeopathy* 2002; 91(4): 197–206.
- 47- McCarthy M. Critics slam draft WHO report on homeopathy. *Lancet* 2005;366(9487):705–706.
- 48- Kleijnen J, Knipschild P, Ter Riet G. Trials of homeopathy. *BMJ* 1991;302(6782):960.