

# 1. EXPERIMENTOS CONTROLADOS

*Hazlo siempre bien. Esto complacerá a algunas personas y dejará atónitas a las restantes.*

MARK TWAIN (ESTADOS UNIDOS, 1835-1910)

## 1.1 Diseño del trabajo de campo sobre la efectividad de la vacuna Salk

Cuando se obtiene un nuevo medicamento en un laboratorio no puede lanzarse al mercado sin antes haber evaluado su efectividad en el mundo real. ¿Cómo se puede diseñar un experimento para contrastar dicha efectividad? La respuesta es que los experimentos de este tipo se basan esencialmente en el llamado método de *comparación*.<sup>1</sup> Éste consiste en administrar el medicamento a un conjunto de sujetos, que constituyen el llamado *grupo en tratamiento*, mientras otro conjunto de sujetos a los cuales no se administra el tratamiento, es empleado como grupo de *control*. Al finalizar el experimento se comparan las reacciones o respuestas de los sujetos de ambos grupos y se extraen las conclusiones pertinentes. El método de comparación debe cumplir además un conjunto de requisitos sin los cuales cualquier experimento carecería de validez. Por ejemplo, los sujetos deben ser asignados al grupo de tratamiento y al de control de forma *aleatoria* y el experimento debe llevarse a cabo con completo *desconocimiento* de los detalles por ambas partes: ni los sujetos ni los investigadores que deberán evaluar las respuestas deben saber qué personas estaban sometidas a tratamiento y qué personas formaban parte del grupo de control. Veamos como fue aplicado este método a un caso real<sup>2</sup>.

La primera epidemia de polio azotó a los Estados Unidos en 1916 y a lo largo de los siguientes cuarenta años cobrándose varios cientos de miles de víctimas, especialmente entre los niños. Hacia 1950 habían sido descubiertas diversas vacunas contra esta enfermedad. Entre ellas, la desarrollada por Jonas Salk, parecía la más prometedora. Fue probada y admitida como segura en experimentos de laboratorio en las cuales había provocado la aparición de anticuerpos contra la polio. Sin embargo, fue necesaria la realización de un trabajo de campo a gran escala para comprobar si la vacuna podía proteger a los niños contra la polio fuera del laboratorio.

En 1954, el Public Health Service de los Estados Unidos decidió organizar este gran experimento empleando como sujetos a miles de niños que pertenecían a los grupos de edad más vulnerables a la enfermedad, es decir, aquellos que cursaban los tres primeros años de escuela primaria. El trabajo de campo se llevó a cabo en distritos escolares seleccionados a través de todo el país entre los lugares en que se pensaba que el riesgo de contraer la enfermedad era muy elevado. Dos millones de niños quedaron involucrados en el experimento de los cuales fueron vacunados quinientos mil. Un millón fue dejado deliberadamente sin vacunar y el medio millón restante rehusó la vacunación.

Como se puede apreciar, este caso real responde al tipo de problema en que resulta adecuado el empleo del método de comparación. Así, los niños que fueron vacunados constituyeron el grupo de tratamiento mientras que los restantes formaron el grupo de control. Al terminar el plazo establecido se efectuó la comparación de los resultados en ambos grupos y se pudo detectar la existencia de diferencias que permitieron probar la efectividad de la vacuna. Sin embargo, como veremos, el desarrollo del experimento no fue tan sencillo.

Lo primero que llama la atención es que en este experimento el grupo en tratamiento y el grupo de control no eran del mismo tamaño. Este hecho carece de importancia debido a que los investigadores compararon los porcentajes en que los niños contrajeron la polio en ambos grupos y no los valores absolutos. En general, la observación de proporciones en lugar de cifras absolutas salva los efectos de las diferencias de tamaños de los grupos analizados.

En segundo lugar surge una cuestión preocupante que atañe a la ética en la medicina: ¿No deberían haber sido vacunados todos los niños? Una respuesta a esta pregunta es que, con medicamentos nuevos, incluso si han sido ampliamente contrastados en el laboratorio, nunca está claro si las ventajas sobrepasan a los riesgos. Siempre se necesita el análisis de los datos proporcionados por un trabajo de campo para averiguar cómo se comporta el tratamiento al ser empleado en el mundo real. Por supuesto, siempre habrá quien piense que la administración de la vacuna a un número muy elevado de niños proporciona una evidencia decisiva, incluso sin el empleo de controles. Por ejemplo, si la incidencia de la polio en 1954 hubiese sido notablemente inferior a la de 1953, habría podido pensarse que quedaba probada la efectividad de la vacuna Salk. Sin embargo, no habría sido así dado que la polio era una epidemia cuya incidencia variaba considerablemente de un año a otro. En 1952 hubo alrededor de 60.000 casos y en 1953 se dieron como mucho la mitad de la cifra anterior. Sin el empleo de controles, una menor incidencia en 1954 podría haber significado una de dos cosas: que la vacuna era efectiva o que no se trataba de un año epidémico. Por tanto, la única forma de saber si la vacuna era efectiva o no era dejar a un conjunto de niños sin vacunar.

El tercer factor a tener en cuenta en el diseño del experimento era el consentimiento paterno a la vacunación: los niños solamente podían ser vacunados bajo

la autorización de sus padres. Frente a este condicionante, el único diseño posible era el siguiente: los niños cuyos padres consintieran en la vacunación formarían el grupo de tratamiento y el resto el grupo de control. Sin embargo, este diseño presentaba problemas ya que se sabía que los padres con niveles de renta más elevados consentirían con mayor frecuencia a las vacunaciones que aquellos que disponían de rentas inferiores, estableciéndose con ello un sesgo importante en el experimento. Este sesgo tendría su origen en el hecho de que los niños procedentes de familias con mayor nivel de renta eran más vulnerables a la polio. Esto que parece paradójico en principio, se explica fácilmente al saber que la polio es una enfermedad relacionada con la higiene. Los niños que vivían en ambientes menos higiénicos tendían a contraer casos de polio más benignos y a una edad más temprana, mientras todavía estaban protegidos por los anticuerpos de sus madres. Después de haber contraído la infección, generaban sus propios anticuerpos que los protegían contra posteriores infecciones de carácter más severo. En cambio, los niños que vivían en ambientes más higiénicos no desarrollaban estos anticuerpos. La lección estadística que hay que extraer de estos últimos comentarios es que hay que evitar todo sesgo, de forma que los grupos a los que se aplique un tratamiento y los que sirvan de control deben ser lo más parecidos posible excepto, precisamente, en lo que se refiere a la aplicación del tratamiento. Es en ese caso que, cualquier diferencia detectada en las respuestas de ambos grupos será debida al tratamiento, en lugar de a cualquier otro factor que pudiera haber incidido en el experimento. Si los dos tipos de grupo difieren entre sí en algún otro factor, los efectos de dicho factor pueden ser confundidos con los del tratamiento. Separar estos efectos puede resultar difícil e incluso imposible. La confusión es una de las mayores fuentes de sesgo en toda investigación.

Debido a todas las consideraciones apuntadas, en el caso del experimento para verificar la efectividad de la vacuna Salk, se presentaron diversas propuestas para diseñar el trabajo de campo. Por un lado, la National Foundation for Infantile Paralysis (NFIP) quería vacunar a todos los niños de segundo curso cuyos padres diesen el permiso pertinente, dejando a los niños de primer y tercer curso como control. Este diseño de la NFIP fue aceptado en varios distritos escolares. Entre otras críticas a este diseño, hay que considerar el hecho de que la polio es una enfermedad contagiosa que se extiende por contacto. Por tanto, la incidencia podría haber sido mayor en el segundo que en el primero y tercer cursos y esto habría sesgado el estudio en contra de la vacuna. También podría haber sucedido que la incidencia hubiese sido menor en el segundo curso, sesgando el estudio a favor de la vacuna. Es más, los niños bajo tratamiento, para el cual era necesario el consentimiento paterno, podían poseer rasgos familiares diferenciales respecto a aquellos que quedarían formando el grupo de control, cuyo consentimiento paterno no era requerido. Con el diseño de la NFIP, el grupo en tratamiento podría incluir a demasiados niños provenientes de familias de rentas elevadas haciendo a este grupo más vulnerable a la polio que el grupo de control. En este caso, existía un sesgo definido en contra de la vacuna.

Estos fallos en el diseño del experimento fueron captados en varios distritos escolares que decidieron adoptar un planteamiento diferente. De esta forma se llegó a un diseño que contemplaba el cumplimiento de los requisitos apuntados al principio de este capítulo. En él, el grupo de control debía ser extraído de la misma población que el grupo que iba a ser vacunado: niños cuyos padres hubiesen consentido en la vacunación. De otra forma, el efecto del entorno familiar habría podido ser confundido con el efecto de la vacuna.

El siguiente paso era decidir la forma en que los niños debían ser asignados a cada grupo. En este caso, podría aplicarse el sentido común para configurar el grupo de control de forma que presentase las mismas características que el de tratamiento en relación con las variables relevantes del estudio, como eran la renta familiar, el estado general de salud de los niños, su personalidad y sus hábitos sociales. Sin embargo, la experiencia demuestra que, a menudo, el sentido común cae en sesgos importantes, siendo mejor el empleo de un procedimiento de elección diseñado cuidadosamente. Para este trabajo de campo, el procedimiento fue equivalente al lanzamiento de una moneda para que cada niño tuviese un 50% de probabilidades de ser asignado al grupo de tratamiento o al de control. Este tipo de procedimiento es objetivo e imparcial ya que las leyes del azar garantizan que, con un número suficiente de sujetos, el grupo de tratamiento y el de control se parecerán mucho en términos de todas las variables importantes del estudio, hayan sido o no identificadas. Cuando se emplea un procedimiento aleatorio para asignar a los sujetos al grupo de tratamiento y al de control, el experimento recibe la denominación de *con control aleatorio*.<sup>3</sup>

Además de las precauciones anteriores, se empleó un *placebo*, es decir, a los niños que formaban el grupo de control les fue aplicada una inyección de sal disuelta en agua. De esta forma, los sujetos que tomaban parte en el experimento no sabían si estaban o no bajo tratamiento. Esta medida era necesaria porque, aunque parezca inverosímil, la gente puede quedar protegida por la fuerza de una idea, es decir, por una autosugestión. Este tipo de fenómeno ha sido observado en muchas ocasiones en pacientes convalecientes de diversas operaciones que sufren dolores post-operatorios de consideración. Dichos pacientes han respondido al tratamiento de sustancias completamente neutras mostrando notables mejorías en sus dolencias.<sup>4</sup>

En la fase final del experimento un grupo de médicos debía decidir si cada sujeto había contraído o no la polio a lo largo del mismo. Dado que muchas formas de polio son de muy difícil diagnóstico, los médicos podrían verse influenciados por el hecho de saber si el niño había sido o no vacunado. Por este motivo, los doctores no fueron informados acerca del grupo al que pertenecía cada niño que examinaron. Puede decirse que gracias a esta medida existía una doble manera de proteger los resultados del experimento: los sujetos no sabían si estaban bajo tratamiento y tampoco lo sabían los que evaluaban su estado de salud. Por eso, el trabajo de campo de la vacuna Salk fue un experimento con control aleatorio y doblemente ciego, que es uno de los mejores diseños que podía aplicarse en este caso.

Los resultados de la experiencia se muestran en la tabla 1, en la que pueden verse las proporciones de casos de polio (para cada 100.000 sujetos), tanto en el grupo de control como en el que estuvo bajo tratamiento. Evidentemente, la proporción de este segundo grupo es mucho menor y constituyó la prueba de la efectividad de la vacuna Salk.

**Tabla 1.** Resultados de las pruebas de la vacuna Salk en 1954. Tamaño de los grupos y proporción de casos de polio por 100.000 en cada grupo. Las cifras están redondeadas.

	Experimento con control aleatorio y doblemente ciego		Estudio de la NFIP	
	Tamaño	Proporción	Tamaño	Proporción
Tratamiento	200.000	28	Segundo curso (Vacunados)	225.000 25
Control	200.000	71	Primer y tercer curso (Control)	725.000 54
No consintieron	350.000	46	Segundo curso (No consintieron)	125.000 44

Fuente: Thomas Francis, Jr., *American Journal of Public Health*, vol. 45 (1955) págs. 1-63.

La tabla 1 también muestra el hecho de que el estudio llevado a cabo por la NFIP presentaba un sesgo en contra de la vacuna. En el experimento controlado de forma aleatoria, la vacuna redujo la proporción de casos de polio de un 71 a un 28 por cada 100.000 niños. En cambio, la reducción aparente que se produce en el estudio de la NFIP (de un 54 a un 25 por cada 100.000 niños), es mucho menor. El principal agente causante del sesgo fue la confusión. El grupo de tratamiento configurado por la NFIP solamente incluía niños cuyos padres habían consentido la vacunación, pero el grupo de control sólo incorporaba niños cuyos padres no habían consentido la vacunación, provocando con ello el que ambos grupos no fueran legítimamente comparables entre sí.

El diseño con control aleatorio y doblemente ciego reducía el sesgo al mínimo y ésta es la razón por la cual debe ser empleado siempre que sea posible. Además de todo lo expuesto, hay que destacar que este diseño posee una ventaja técnica importante. Para hacerla patente vamos a hacer por un momento de abogados del diablo. Supongamos que la vacuna Salk no hubiese tenido en realidad ninguna efectividad. En ese hipotético caso, la diferencia entre la proporción de casos de polio del grupo de tratamiento y del grupo de control se habría debido únicamente al azar, pero ¿cómo explicar entonces su magnitud?

En el diseño de la NFIP, los resultados están afectados por un conjunto de factores que (desde el punto de vista de los investigadores) son aleatorios: la voluntad de las

familias, el que los niños estén cursando segundo curso, y así sucesivamente. Debido a esta circunstancia, los investigadores no disponen de suficiente información para evaluar los resultados desde la óptica probabilística y, por tanto, no pueden calcular qué parte de las diferencias en las proporciones de casos de polio son debidas a estos factores accidentales. En cambio, con un experimento con control aleatorio el azar interviene de una forma simple y planificada en el acto de asignar los niños a cada uno de los grupos.

Siguiendo como abogado del diablo, nuestra hipótesis es que la vacuna no tiene efecto. Según esta hipótesis, muchos niños están destinados a contraer la polio, de manera que, asignarlos al grupo de tratamiento o de control es indiferente. Cada niño tiene una probabilidad del 50% de ser enviado a uno u otro grupo dependiendo del resultado del lanzamiento de una moneda. En estas circunstancias, cada caso de polio tiene una probabilidad del 50% de aparecer en el grupo de tratamiento o en el de control. Por tanto, el número de casos de polio que se daría en los dos grupos debería ser muy parecido. Cualquier diferencia, según esa hipótesis, sería debida a la variabilidad propia del lanzamiento de una moneda. Se puede calcular la probabilidad de que una diferencia como la observada sea debida a esta variabilidad. El cálculo lo haremos en el capítulo 27 y la probabilidad resulta ser extraordinariamente pequeña, del orden de uno en mil millones.

## 1.2. La derivación porta-cava

Una gran parte de los pacientes con cirrosis hepática presentan hemorragias internas que pueden provocarles la muerte. Un tratamiento quirúrgico que trata de evitar estas hemorragias consiste desviar el flujo sanguíneo a través de una porta-cava. La operación para crear esta derivación es larga y complicada. Ante esta situación es lícito preguntarse si los beneficios de la operación superan los riesgos. Para responder a ello se han llevado a cabo unos 50 estudios que evalúan los efectos de esta cirugía.<sup>5</sup> Los resultados de dichos estudios quedan resumidos en la tabla 2.

Como puede apreciarse en la tabla, se llevaron a cabo 32 estudios sin elementos de control, de los cuales el 75% proporcionaron resultados notablemente optimistas. La conclusión de este tipo de experimentos fue, como es lógico, que los beneficios superaban ampliamente los riesgos. Por otro lado, se llevaron a cabo 15 estudios con elementos de control, en los cuales, la asignación al grupo de tratamiento y al de control no se realizó de forma aleatoria. Respecto a los resultados de este segundo grupo de estudios hay que decir que sólo un 67% fueron notablemente entusiastas acerca de la aplicación de esta cirugía. Finalmente, cuatro estudios contemplaron elementos de control con asignación aleatoria y dieron como resultado un escaso o nulo valor a esta técnica quirúrgica. Por tanto, los estudios mal diseñados exageraron el valor de esta arriesgada operación.

**Tabla 2.** Análisis de 51 estudios sobre la derivación porta-cava. Los estudios bien diseñados mostraron que la aplicación de esta técnica quirúrgica tenía un escaso o nulo valor. Los estudios mal diseñados exageraron su valor.

Diseño	Grado de optimismo		
	Notable	Moderado	Nulo
Sin controles	24	7	1
Con controles no aleatorios	10	3	2
Con controles aleatorios	0	1	3

Fuente: Grace, Muench y Chalmers, *Journal of Gastroenterology*, vol. 50, (1966) págs. 646-91.

La explicación de todo ello reside en el hecho de que en un experimento sin elementos de control o en un experimento en el cual los sujetos son asignados a los grupos de tratamiento y de control como resultado de un dictamen médico, existe una tendencia natural a tratar únicamente a los pacientes que presentan un estado físico relativamente bueno, y es indudable que este hecho sesga el estudio a favor del tratamiento. En los tres tipos de estudio de la tabla 2, se da la circunstancia de que alrededor del 60% de los pacientes que forman parte del grupo que se sometió al tratamiento estaban vivos tres años después de la operación. En los experimentos controlados de forma aleatoria, el porcentaje de personas del grupo de control que sobrevivieron al experimento tres años fue también cercano al 60%, pero sólo un 45% de estas personas sobrevivieron tres años en el caso de los experimentos no aleatorios. Estos resultados muestran el sesgo en que incurren este tipo de diseños: los pacientes en peor estado de salud fueron empleados como control, mientras que los que estaban en mejores condiciones físicas fueron colocados en el grupo de tratamiento. En definitiva, lo que ocurría es que el grupo de control no era comparable al grupo sometido a tratamiento.

### 1.3 Controles históricos

Los experimentos con control aleatorio presentan el inconveniente de ser laboriosos y complicados. Por este motivo, con frecuencia, los médicos emplean otros tipos de diseño. Por ejemplo, en algunas ocasiones prueban un nuevo tratamiento en un grupo de pacientes y comparan los resultados con un conjunto de sujetos denominados "controles históricos", que consiste en un grupo de pacientes que recibió un tratamiento más antiguo en el pasado. El problema que presenta este tipo de diseño es que el grupo actualmente en tratamiento puede presentar importantes diferencias relativas al tratamiento respecto al grupo de control histórico. Por ejemplo, algunos trabajos de campo sobre la porta-cava realizados con escaso o nulo control, se llevaron a cabo mediante la comparación con controles históricos. (También se llevaron a cabo algunos estudios con controles contemporáneos, pero en ellos, la asignación

de los sujetos al grupo de control no fue realizada de forma aleatoria.) De todo lo expuesto hasta el momento se desprende que el diseño, en general, es un tema muy importante al que hay que prestar gran atención para poder obtener resultados que sean estadísticamente fiables. Por eso, vale la pena profundizar un poco más con algún otro ejemplo.

En la misma línea, la técnica quirúrgica del "bypass" es una operación muy cara y frecuentemente aplicada en casos de lesiones de las arterias coronarias. Chalmers y sus asociados pudieron acceder a información proveniente de veintinueve trabajos de campo sobre esta operación y que queda resumida en la primera línea de la tabla 3. Como se puede observar, ocho de los estudios se llevaron a cabo mediante trabajos de campo con control aleatorio, siete de los cuales dieron resultados bastante negativos sobre las ventajas de esta operación. Por contra, veintiuno de los trabajos fueron realizados con controles históricos, lográndose resultados positivos en dieciséis de los mismos. Por tanto, los estudios peor diseñados resultaron ser mucho más entusiastas que los mejor diseñados respecto a la validez de la operación. Las restantes líneas de la tabla pueden ser leídas de la misma forma y conducen a idénticas conclusiones.

**Tabla 3.** Análisis de un conjunto de estudios sobre cinco terapias evaluadas a partir de trabajos de campo diseñados con controles aleatorios y de trabajos de campo en que se emplean controles históricos. Las conclusiones de dichos trabajos son evaluadas como positivas respecto a la terapia cuando aparece el signo (+) y como negativas cuando aparece el signo (-).

Terapia	Controlado de forma aleatoria		Controles históricos	
	+	-	+	-
Cirugía coronaria bypass	1	7	16	5
Anticoagulantes	1	9	5	1
5-FU	0	5	2	0
BCG	2	2	4	0
DES	0	3	5	0

Nota: Los anticoagulantes se emplean para el tratamiento de los infartos de corazón; el 5-FU se emplea en quimioterapia para el cáncer de colon; el BCG se emplea en el tratamiento de los melanomas y el DES para prevenir los abortos. Fuente: Chalmers, Sacks y Smith, *American Journal of Medicine*, vol. 72 (1982) págs. 233-40.

En seis de los experimentos con control aleatorio sobre la técnica del bypass y en nueve de los estudios realizados con el empleo de controles históricos, se elaboró un informe sobre los porcentajes de supervivencia una vez transcurridos tres años desde la operación, tanto para los sujetos del grupo de tratamiento como para los del grupo de control. Estos resultados, resumidos en la tabla 4, muestran que la supervivencia de los sujetos analizados en el experimento con control aleatorio fue muy similar en



ambos grupos. Ésta fue la causa de que los investigadores no mostrasen un excesivo entusiasmo hacia esta técnica quirúrgica, ya que no servía para salvar vidas.

Si observamos los resultados de los estudios que emplearon controles históricos constatamos que la supervivencia en el grupo de pacientes operados es muy similar a la del otro tipo de experimento. En cambio, los porcentajes de supervivencia de los pacientes del grupo de control son mucho más bajos. La razón radica en el estado de salud de los mismos, que no era suficientemente buena como para aconsejar la aplicación de la cirugía, mientras que sí lo era en los pacientes a los que se aplicó el tratamiento. Los trabajos de campo que emplearon controles históricos están sesgados a favor de la cirugía mientras que los experimentos aleatorios evitaron este tipo de sesgo.

Volviendo a los resultados de la tabla 3, la última línea es la que puede provocar mayores discusiones. La DES (diethylstilbestrol) es una hormona artificial que se emplea en la prevención del aborto espontáneo. Chalmers y sus colegas hallaron los resultados de ocho trabajos de campo dedicados a su evaluación. Tres de ellos eran con control aleatorio y todos presentaban conclusiones negativas: según ellos, esta droga no era efectiva. En cambio, cinco de los estudios habían sido diseñados con el concurso de controles históricos, mostrando en todos los casos resultados positivos. Estos estudios, mal diseñados, presentaban un sesgo a favor de la terapia.

Se prestó muy poca atención a los experimentos con control aleatorio. Incluso a finales de los años sesenta se estaba administrando esta hormona a 50.000 mujeres por año. Todo ello constituyó una gran tragedia médica, tal y como fue demostrado por estudios posteriores. Si se administraba durante el embarazo, el DES podía tener un efecto desastroso veinte años después, causando en las hijas el desarrollo de una extraña forma de cáncer de vagina. En 1971 se descartó el uso del DES en el caso de mujeres embarazadas.<sup>6</sup>

**Tabla 4.** Porcentajes de supervivientes después del transcurso de tres años para pacientes operados y del grupo de control en los trabajos de campo sobre la técnica del bypass. Los experimentos con control difieren de los llevados a cabo con el empleo de controles históricos.

	Controles aleatorios	Controles históricos
Cirugía	87,6%	90,9%
Control	83,2%	71,1%

Nota: Se realizaron seis experimentos controlados de forma aleatoria con el concurso de 9.290 pacientes y nueve estudios con controles históricos con un total de 18.861 pacientes.  
Fuente: Véase tabla 3.

### 1.4 Resumen

1. En estadística se emplea a menudo el método comparativo. Si se desea averiguar el efecto que produce un *tratamiento* (como la Salk) sobre un resultado (como contraer la polio), se comparan los resultados sobre un *grupo en tratamiento* con los obtenidos sobre un *grupo de control*. Habitualmente es difícil juzgar el efecto de un tratamiento de forma apropiada sin efectuar algún tipo de comparación.
2. Si la similitud entre el grupo en tratamiento y el de control es muy estrecha, excepto por lo que se refiere al tratamiento, la diferencia entre las respuestas de ambos grupos puede considerarse debida a los efectos del tratamiento.
3. En cambio, si el grupo en tratamiento difiere del grupo de control en otros elementos, los efectos de éstos se confunden con el efecto del tratamiento.
4. Para asegurarse de que ambos grupos son iguales, hay que asignar los sujetos a cada uno de los grupos de forma aleatoria. Esta técnica se emplea en los experimentos con control aleatorio.
5. Siempre que sea posible, al grupo de control le será administrado un *placebo* sin efectos que no se pueda distinguir del tratamiento. La respuesta obtenida tiene que deberse al tratamiento y no a la idea de haber recibido el tratamiento.
6. En los experimentos doblemente ciegos, los sujetos no saben si pertenecen al grupo de tratamiento o al de control y tampoco lo saben las personas que deban evaluar las respuestas. Esta forma de actuar constituye una protección contra un posible sesgo, tanto en las respuestas como en su evaluación.

Tabla 4. Evolución de supervivencia después del transcurso de tres años para pacientes operados y del grupo de control en los trabajos de campo sobre la técnica del papiros. Los experimentos con control difieren de los llevados a cabo con el empleo de controles históricos.

	Controles históricos	Controles históricos
Control	87,5%	71,1%
Operados	80,9%	71,1%

Note: Se realizaron seis experimentos controlados de forma aleatoria con el consumo de 2.100 papiros y nueve estudios con controles históricos con un total de 12.861 pacientes. Fuente: Véase tabla 3.