PubliCE, (PubliCE) 2004

Meta-analysis

Los Estudios Meta-Analíticos no Respaldan la Realización de Series Múltiples o el Alto Volumen en el Entrenamiento de la Fuerza

Richard A Winett[1]

[1]*Center for Research in Health Behavior, Department of Psychology, Virginia Tech, Blacksburg, VA, 24061-0436, USA.*

RESUMEN

Cuatro estudios meta analíticos publicados recientemente afirman que sus resultados muestran que los protocolos de entrenamiento de la fuerza con múltiples series (alto volumen) son superiores a aquellos que utilizan una serie única para cada ejercicio (bajo volumen) para producir ganancias de fuerza en individuos con experiencia en el entrenamiento. Esta crítica examina el marco contextual, la lógica, los procedimientos, los análisis estadísticos, los resultados y las interpretaciones de los cuatro estudios meta-analíticos y muestra que estos estudios no siguieron muchas de las normas reconocidas para la realización de un meta-análisis. Hubo muy poco respaldo para cualquiera de las afirmaciones o conclusiones. De hecho, esta crítica sugiere que los protocolos simples, con series únicas, eficientes en cuanto al tiempo y de bajo volumen parecen ser tan efectivos como los protocolos con series múltiples y alto volumen para incrementar la fuerza muscular, sin tener en cuenta los objetivos o el nivel de entrenamiento.

Palabras Clave: levantamiento de pesas, volumen, intensidad, adaptación

INTRODUCCION

Desde hace tiempo se sostiene un gran debate acerca de si la realización de series múltiples en cada ejercicio o un gran volumen de ejercicios por grupo muscular resulta en mayores ganancias de fuerza en comparación a una única serie por ejercicio. Existe poca evidencia que respalda la mayor eficacia de las series múltiples (1), y una reciente reexaminación de una estudio de hace 40 años, el cual fue quizás la génesis de la creencia en el entrenamiento de la fuerza con series múltiples, reveló que hay muy poca evidencia que respalde la utilización de series múltiples (2). Sin embargo, el Colegio Americano de Medicina del Deporte (ASCM) en su declaración de posición para el entrenamiento de la fuerza recomienda series múltiples para sujetos con experiencia en el entrenamiento y para atletas competitivos (3).

Las inconsistencias en los resultados y en la interpretación de los datos entre los diferentes estudios pueden a menudo deberse a problemas en la fortaleza estadística de algunas o todas las investigaciones en cuestión. En un esfuerzo por superar las limitaciones impuestas por la fortaleza estadística en las investigaciones previas, los investigadores y académicos a menudo utilizan el meta-análisis. Un meta-análisis es una vía cuantitativa para integrar los resultados de estudios empíricos de un solo campo de investigación, tal como los estudios acerca del entrenamiento de la fuerza. El propósito del meta-análisis es comparar los resultados producidos en los diferentes estudios por medio de diferentes

tratamientos o protocolos tales como los resultados obtenidos en la fuerza a partir de entrenamientos con series únicas o con series múltiples. Con el objeto de ser incluido en un meta-análisis en un campo determinado, cada estudio necesita tener una métrica común. Por ejemplo, en un meta-análisis acerca del entrenamiento de la fuerza, cada estudio deberá tener una medición de la fuerza como resultado del protocolo de entrenamiento.

Cada uno de los estudios incluidos contribuirá al tamaño del efecto del meta-análisis. El tamaño del efecto es una medida estadística estándar que representa el grado o cantidad de "resultados", "impactos" o "efectos", y se cuantifica como la diferencia media dividida por el desvío estándar. Un tamaño del efecto de 0.00 a 0.32 es considerado "pequeño", un tamaño del efecto de 0.33 a 0.55 es considerado "mediano" y un tamaño del efecto de 0.56 a 1.20 es considerado "grande" (4). Hay cierta relevancia para dicha categorización y cuanto mayor es el tamaño del efecto mejor.

El tamaño del efecto puede calcularse a partir de cada estudio sobre la fuerza muscular producida como resultado de un tratamiento o protocolo, tal como el entrenamiento con series únicas o el entrenamiento con series múltiples. En un meta-análisis, por ejemplo, la media total para todos los tamaños del efecto a partir de los estudios en donde se utilizaron protocolos con series múltiples, puede compararse con la media total de todos los tamaños del efecto a partir de los estudios en donde se utilizaron protocolos con series únicas, y de esta manera determinar si existe alguna diferencia estadísticamente significativa entre ellos. Sin embargo, el meta-análisis es viable no solamente si se siguen normas establecidas para la metodología, la estadística y la interpretación. Las conclusiones extraídas de un meta-análisis serán buenas siempre y cuando sean buenos los criterios utilizados para seleccionar las investigaciones a utilizar en el análisis.

Cuatro estudios meta-analíticos recientes (5-8) afirman que sus resultados proporcionan evidencia convincente de que cuando los datos de los estudios sobre entrenamiento de la fuerza son agregados apropiadamente, el tamaño del efecto para los resultados de la fuerza son considerablemente mayores cuando se utilizan series múltiples o un gran volumen de ejercicios por grupo muscular en comparación con una serie única por ejercicio. Se ha hipotetizado que estos cuatro estudios meta-analíticos contienen errores y que ninguno de estos meta-análisis proporciona lógica, teórica o estadísticamente, alguna evidencia real de la superioridad del entrenamiento con series múltiples por ejercicio o con un gran volumen de ejercicios por grupo muscular. También se ha hipotetizado que ninguno proporciona evidencias claras acerca de la importancia de algún parámetro específico del entrenamiento o del protocolo ya que los cuatro meta-análisis (5-8), en general, no han seguido varias de las normas clave para la realización de un meta análisis propuestas por Lipsey y Wilson (9), incluyendo:

1. La separación de los estudios diferenciando rigurosamente los diseños de investigación (i.e., estudios con diseños aleatorios y grupo control versus estudios con diseño no aleatorio y otros estudios);
2. Evaluación de la homogeneidad para determinar si hay mayor variación en algún grupo de los estudios categorizados (e.g., cuatro series vs. dos series por grupo muscular) que la variación esperada a partir del error de muestreo (indicando que estudios no deberían ser agrupados y categorizados);
3. Utilización de un tamaño del efecto medio en un estudio y no efectos del tamaño múltiples no independientes;
4. Identificar estudios con resultados atípicos y tamaños del efecto tanto moderadores como excluyentes;
5. Reconocimiento de grandes variaciones del tamaño del efecto dentro de un grupo y no enfocarse solamente en los valores medios del tamaño del efecto;
6. Reportar los resultados utilizando intervalos de confianza para expresar un rango probable de resultados en una población; y,
7. Tener una guía teórica de las hipótesis estudiadas, con los resultados conformando la teoría o la posibilidad biológica.

Los meta-análisis se están utilizando ampliamente en muchos campos y los investigadores han señalado las limitaciones y temas relacionados al meta-análisis. Esto incluye la necesidad de examinar críticamente cada estudio incluido en el meta-análisis para asegurar la calidad de cada estudio y la precisión de las conclusiones (10); evaluando y reportando las fuentes de sesgo en los estudios que son incluidos en el meta-análisis y reportando los sesgos sistemáticos que pueden existir entre los estudios (11, 12); entendiendo y reportando la variación de los efectos de tratamiento (13-15); incluyendo moderadores y mediadores de los efectos de tratamiento (16); teniendo precaución cuando se reportan meta-análisis realizados con pocos estudios y efectos de tamaño (17) o donde los estudios en el meta análisis tienen poca diferencia (15), y no reducir la complejidad de un campo y la efectividad del tratamiento al "santo grial" de la media para el tamaño del efecto (12, 18).

El propósito de este documento es revisar cada uno de éstos meta-análisis en detalle a la vez que se realiza una crítica a cada estudio.

META-ANALISIS

Rhea y cols. (5)

Rhea y cols. (5) realizaron un meta-análisis en el que afirmaron incluir todos los estudios publicados y no publicados que tuvieran intervenciones en el entrenamiento de la fuerza, aunque con diferentes diseños experimentales, y donde hubiera una medición de la fuerza pre y post-entrenamiento. Una examinación de su extensa lista de referencias indicó que al menos 26 estudios que cubrían estos criterios no fueron incluidos en su análisis. Estos estudios son citados en una crítica reciente de la declaración de posición del ACSM (19); 24 de los 26 estudios incluidos mostraron diferencias no significativas entre series únicas y series múltiples por ejercicio (19; ver p. 17, Tabla 4). La exclusión de los estudios específicos puede crear sesgos en los resultados del meta-análisis, particularmente si los resultados reportados están consistentemente en contra de las conclusiones del meta-análisis.

Para cada estudio en el meta-análisis; Rhea y cols. (5) reportaron efectos del tamaño (ES) para la intensidad (% de 1RM), frecuencia (número de días por semana en que los participantes entrenaron un grupo muscular en particular) y para el volumen (número de series por grupo muscular durante cada sesión), en sujetos previamente desentrenados y con experiencia (>1 año) en el entrenamiento de la fuerza. Rhea y cols. (5) reportaron diferencias estadísticamente significativas en los ES entre los sujetos previamente desentrenados y los sujetos con experiencia, pero no quedo claro si esta diferencia significativa eran para la intensidad, la frecuencia o el volumen de entrenamiento. La utilización de series por grupo muscular en comparación con la utilización de series por ejercicio es el punto de partida para entender como se han analizado y discutido estos estudios. Sin embargo, el artículo no contiene información acerca de cómo se clasificaron los ejercicios por grupo muscular y solo una persona codificó las series por grupo muscular, por lo tanto, la estadística disponible no es confiable. Su codificación no es directa. Por ejemplo, no se especifica si el ejercicio de sentadilla es categorizado como un ejercicio que afecta a los principales grupos musculares involucrados en la realización de la sentadilla (glúteo, cuadriceps, isquiotibiales, gastronemio y los erectores de la columna) o como un ejercicio que afecta solamente al cuadriceps.

Para cada estudio, y donde fuera posible para cada variable, el tamaño del efecto se generó en base a la diferencia entre los valores de la fuerza pre y post-entrenamiento dividido por el desvío estándar del valor de la fuerza pre-entrenamiento indicando de este modo la efectividad de una variable o tratamiento dado. Se generaron cientos de tamaños del efecto. El resultado es que algunos estudios contribuyeron de forma inapropiada con muchos tamaños del efecto mientras que otros estudios contribuyeron con unos pocos, lo cual crea sesgos potenciales (9).

Debido a que los estudios tenían diferentes números de participantes, Rhea y cols. (5) utilizaron un factor de ponderación para el número de participantes. Para que una variable fuera considerada, era necesario que hubiera al menos 10 tamaños del efecto a partir de todos los estudios incluidos en el meta-análisis. También fue posible observar los efectos en sujetos previamente desentrenados (<1 año de entrenamiento) y en sujetos con experiencia (>1 año de entrenamiento). Rhea y cols. (5) calcularon entre los estudios, el tamaño del efecto medio total para cada variable para sujetos entrenados y desentrenados.

Una fortaleza de la metodología utilizada por Rhea y cols. (5) es la inclusión de estudios publicados (aunque se excluyeron algunos estudios publicados; ver mas arriba) y de estudios no publicados, para limitar el “efecto de exclusión de archivos”. Efecto de exclusión de archivos significa que los estudios sin resultados estadísticamente significativos son relegados a los archivos excluidos y no son publicados. Si estos estudios no son incluidos, el resultado es que los tamaños del efecto totales aumentan. Una debilidad de la metodología utilizada por estos autores fue no darle mayor peso a los estudios que utilizaron mejores metodologías, como aquellos estudios que realizaron la aleatorización de las condiciones. Otro problema fue que sus presentaciones de los datos no examinaron independientemente los efectos de una variable, e.g., series. La variable series/grupo muscular en cualquier estudio y desde un punto de vista estadístico se confunde con la intensidad y la frecuencia. Se deberían haber examinado los efectos conjuntos de estas variables o se deberían haber mantenido otras variables constantes (intensidad, frecuencia) para examinar el efecto de una variable (series). La codificación debería haberse realizado para el porcentaje de 1RM para permitir la utilización de esta variable como moderadora en el análisis.

Como se esperaba, los tamaños del efecto para los sujetos previamente desentrenados fueron mayores que los tamaños del efecto para los sujetos con experiencia en el entrenamiento de la fuerza, simplemente debido a que el incremento en la fuerza se produce en mayor medida y más rápidamente en sujetos desentrenados. En forma global, el meta-análisis reportado por Rhea y cols. (5), presentado principalmente con tablas y figuras y análisis limitados, sugiere que para los sujetos experimentados los parámetros más efectivos para producir ganancias de fuerza son el entrenamiento al 80% de 1RM (alrededor de 8 repeticiones, utilizando una duración convencional por repetición, ver más adelante), entrenar los grupos musculares dos veces por semana (tres veces fue menos efectivo), y realizar cuatro series por grupo muscular. De acuerdo con Rhea y cols. (5), los sujetos previamente desentrenados, aparentemente obtienen mejores resultados

entrenando al 60% de 1RM, tres veces por semana, y con cuatro series por grupo muscular.

Las conclusiones se basaron simplemente en el patrón de los resultados y en “la tendencia de los gráficos” (representación gráfica de los datos), sin reportar otros tipos de análisis. Dicha tendencia de graficación puede llevar a conclusiones que no pueden ser respaldadas. Por ejemplo, mientras que el tamaño del efecto para los sujetos experimentados utilizando cuatro series por grupo muscular (1.17±0.81) fue mayor que con la utilización de dos series por grupo muscular (0.92±0.52), yo realicé una prueba t utilizando los datos del estudio y hallé que esta diferencia no era significativa ($t<1$). De esta manera, los datos reales del meta-análisis indicaron que no hay diferencias significativas entre entrenar con dos o cuatro series por grupo muscular.

Una inspección de cerca de los datos presentados en el meta-análisis revela algunos puntos adicionales. En la Tabla 1 (pag. 9) se presentan datos sobre el entrenamiento a diferentes porcentajes de 1 RM, junto con el tamaño del efecto para los incrementos de fuerza que requieren de una explicación. Los tamaños del efecto de los estudios con sujetos previamente entrenados fueron de 0.70±0.65 para el 70% 1RM, 0.74±0.99 para el 75% 1RM, 1.80±1.30 para el 80% 1RM y 0.65±0.77 para el 85% 1RM. Los desvíos estándar relativamente grandes a un porcentaje dado de 1RM indican que los estudios y sus resultados individuales dentro del porcentaje de 1RM difieren unos de otros en cierta forma significativa. Esto sugiere que debió haber alguna variable de moderación importante que necesitó ser señalada o analizada por Rhea y cols. (5), y sin embargo fue descuidada.

Tampoco hay actualmente una teoría o un mecanismo fisiológico identificable que explique porque el 75% y el 85% de 1RM constituyen estímulos efectivos, pero el 80% es extremadamente efectivo. Lo que presumiblemente signifique esto es que si para un ejercicio la 1RM es 100kg entonces la realización de un número específico de repeticiones hasta la fatiga con 75 u 85kg en un ejercicio específico sea efectivo, pero realizar repeticiones hasta la fatiga con 80kg sea mucho más efectivo. No existe mecanismo fisiológico que explique porque este podría ser el caso. Rhea y cols. (5) fallaron en argumentar este punto.

Las investigaciones muestran que hay un rango considerable de repeticiones que pueden ser realizadas a un RM dada por diferentes individuos y con diferentes tipos de ejercicios para el entrenamiento de la fuerza (19). Por ejemplo, en un estudio reciente, Chromiak y cols. (20) hallaron que cuando se utilizó el 85% de 1RM los sujetos fueron capaces de realizar una media de 4.5±1.40 y 4.7±1.70 repeticiones en el ejercicio de press en banca, pero pudieron realizar 8.8±4.5 y 10.8±6.1 repeticiones en el ejercicio de prensa de piernas pre y post-entrenamiento, respectivamente. Los diferentes números de repeticiones y los diferentes desvíos estándar en los diferentes ejercicios al mismo porcentaje de 1RM llaman ponen en duda, como se señala en otros estudios (19), a los modelos de entrenamiento tales como los prescritos en la Declaración de Posición del ACSM (3) que utilizan un porcentaje específico de 1RM para un número específico de repeticiones. Si los datos del 80% en 1 RM en el meta-análisis de Rhea y cols. (5) representan algún artificio, entonces el entrenamiento a cualquier RM desde el 70% al 85% de 1RM es igualmente efectivo. El patrón de los resultados reportados por Rhea y cols. (5) no tiene sentido teórico y carece de posibilidad fisiológica.

Rhea y cols. (5) señalaron que los resultados se su meta-análisis solamente sugieren los parámetros para los protocolos más óptimos para el desarrollo de la fuerza. En base al nivel de entrenamiento y a los objetivos personales además de otras actividades físicas, una persona podría elegir realizar un bajo volumen de entrenamiento. Por ejemplo, el número de series por grupo muscular o la frecuencia de entrenamiento podrían reducirse si la persona ha alcanzado el punto en donde las adaptaciones han disminuido en lo que se refiere a la fuerza o a la masa muscular. Sin embargo, no hay datos presentados por Rhea y cols. (5) que respalden cualquier protocolo particular de entrenamiento.

En resumen, los resultados del meta-análisis realizado por Rhea y cols. (5) son cuestionables debido a que excluyeron ciertos estudios publicados, utilizaron muchos tamaño del efecto no independientes, se enfocaron en las medias de los tamaños del efecto y le prestaron menor atención a la gran variación dentro de una categoría de entrenamiento dada, y no pudieron explicar el patrón de los resultados con alguna teoría o mecanismo fisiológico.

Wolfe y cols. (6)

Wolfe y cols. (6) afirmaron que su meta-análisis mostraba la superioridad de utilizar series múltiples por ejercicio para incrementar la fuerza. Estos autores realizaron un meta-análisis de los estudios publicados que investigaron el impacto de series únicas vs. series múltiples de entrenamiento, que cumplían al menos con los criterios metodológicos mínimos, y a partir de los cuales se pudiera extraer un tamaño del efecto. Wolfe y cols. (6) controlaron algunas de las variables importantes dentro de la muestra de estudios seleccionados tales como, sesgo de publicación, calidad del diseño de investigación, número de sujetos en cada estudio, número de tamaño del efecto generados, y nivel de valores atípicos. Sin embargo, luego de examinar el artículo de Wolfe y cols. (6), se puede notar por su narrativa que los estudios identificados como atípicos no fueron excluidos.

Wolfe y cols. (6) enfrentaron el problema de los múltiples tamaños del efecto no independientes, incluyendo en primer

lugar todos los tamaños del efecto extraídos de un estudio y luego incluyendo la media del tamaño del efecto de los estudios con múltiples tamaños del efecto. No queda claro como el enfoque utilizado por Wolfe y cols. (6) corrigió el problema.

Sorprendentemente, luego de cuatro décadas de investigación en esta área, Wolfe y cols. (6) observaron que solamente 16 estudios cumplían aceptablemente con los criterios para ser incluidos en su análisis. Sin embargo, Wolfe y cols. (6) en la Tabla 1 (pag. 11) indicaron que habían incluido 20 estudios en su meta análisis. *Solamente seis de los estudios incluían sujetos con experiencia previa en el entrenamiento de la fuerza*. Un problema es simplemente el muy limitado número de estudios a partir del cual extraer conclusiones. La utilización de un pequeño número de estudios en un meta-análisis puede resultar en un error de muestreo de segundo orden; este es un error de muestreo a nivel del meta-análisis (9).

Wolfe y cols. (6) indicaron que el tamaño del efecto para los sujetos entrenados fue significativamente mayor como resultado del entrenamiento con series múltiples (0.70) en comparación con series únicas (0.29), pero no fue significativamente diferente con respecto a los sujetos desentrenados (TE=1.73 y 1.69, series múltiples y series únicas, respectivamente). Sin embargo, hay otros problemas con el meta-análisis. Wolfe y cols. (6) no definieron el estado de entrenado para un sujeto. Wolfe y cols. (6) señalaron en la Tabla 3 que un estudio realizado por Kraemer y cols. (22) involucró a sujetos entrenados. Sin embargo, como se indicó en una crítica previa (19), los sujetos eran jugadoras de tenis de nivel universitario previamente desentrenadas. Wolfe y cols. (6) citaron el artículo de Kraemer y cols. (22) como una publicación de *Medicine and Science in Sports and Exercise*, cuando de hecho es una publicación del *American Journal of Sports Medicine*. Wolfe y cols. (6) también afirmaron que con solo una excepción, cada estudio incluido en su meta-análisis que involucraba a sujetos entrenados reportaba la superioridad de las series múltiples (pag. 44). En una sección posterior, los estudios señalados por Wolfe y cols. (6) son delineados y se demuestra que esta afirmación es incorrecta. Sin embargo, los que discutiblemente pueden ser pequeños errores cuestionan la precisión de los investigadores que están afrontando un desafío mucho más complicado como la realización de un meta-análisis que produce 103 tamaños del efecto.

Los resultados de Wolfe y cols. (6) incluyen seis estudios que reportaron hallazgos para sujetos con experiencia. Como se señaló anteriormente, los resultados mostraron que el tamaño del efecto para los protocolos con una serie única fue de 0.29, para los protocolos con series múltiples, el tamaño del efecto fue de 0.70 y hubo diferencias significativas entre estos tamaños del efecto. Los datos aparentemente muestran que la realización de series múltiples por ejercicio puede ser mucho más efectivo para los sujetos experimentados. Sin embargo, los estudios con sujetos experimentados incluidos en este meta-análisis requieren una mayor examinación (6).

Una serie de experimentos realizados por Kraemer (21) contribuyó con 10 tamaños del efecto debido a que sus datos fueron extraídos de tres experimentos (#2-4) en un solo artículo. En 1997 Kraemer reportó experimentos que había realizado 15 años antes como entrenador de Fútbol Americano. En el experimento #2 Kraemer (21) reportó que los jugadores que entrenaron con series únicas mostraron menores ganancias, expresadas tanto como tamaño del efecto o como porcentaje de cambio, en los ejercicios de press de banca (0.06; 4%) y prensa de piernas (0.13; 4%). Los jugadores que entrenaron con series múltiples mostraron mayores ganancias en los ejercicios de press de banca (0.13; 13%) y en prensa de piernas (0.77; 19%). Para el experimento #3, las ganancias de fuerza en press de banca con series únicas (0.07; 3%) fueron considerablemente menores que con series múltiples (0.24; 11%); similares resultados se reportaron para el envión colgado. Para el experimento #4, el incremento en la fuerza en la prensa de piernas (el único ejercicio reportado con 1RM) con series únicas (0.24; 6%) también fue considerablemente menor que con series múltiples (0.64; 18%). Los grandes incrementos en la fuerza utilizando series múltiples para este grupo de atletas con un alto nivel de fuerza son inusuales, o incluso cuestionables; y los resultados deberían ser tratados como atípicos (9).

Un estudio mostró diferencias entre el entrenamiento con series únicas y con series múltiples (22). Sin embargo como se señaló anteriormente, el nivel de entrenamiento de los sujetos al comienzo del estudio fue designado de manera incorrecta por Wolfe y cols. como sujetos experimentados. Una examinación del protocolo del estudio indicó que las mujeres participantes necesitaron un periodo de tiempo para familiarizarse con el entrenamiento de la fuerza.

Otros dos estudios (25, 26) con sujetos experimentados en el análisis son de interés. En un estudio de 13 semanas de duración en donde los participantes entrenaron tres veces por semana, Hass y cols. (25) compararon la realización de una y tres series en un circuito de entrenamiento que consistía de nueve ejercicios para el entrenamiento de la fuerza. Los datos fueron presentados como resultados para cinco ejercicios. Yo he realizado el cálculo de los tamaños del efecto para las dos condiciones que mostraron casi los mismos resultados excepto para el ejercicio de flexión de rodillas (0.33 con series únicas y 1.0 con series múltiples). Las ganancias de fuerza para los sujetos con experiencia en el entrenamiento fueron de aproximadamente 10% sin tener en cuenta el protocolo utilizado (una o tres series). Hubo un incremento significativo de la fuerza en todos los ejercicios como resultado de la realización de una serie o de tres series en cada ejercicio, sin diferencias significativas entre los resultados de grupos que realizaron una y tres series.

Ostrowski y cols. (26) compararon protocolos que implicaban la realización de una, dos o cuatro series por ejercicio. Cada

ejercicio se realizó una vez por semana y los participantes entrenaron cuatro veces por semana utilizando rutinas partidas. Los participantes realizaron tres ejercicios por grupo muscular por lo cual algunos sujetos realizaron tres series por grupo muscular por semana (grupo de 1 serie) mientras que otros realizaron hasta 12 series por grupo muscular por semana. Al final del estudio, no hubo diferencias significativas de los resultados entre los grupos. Las ganancias de fuerza fueron de alrededor del 6%, más o menos lo que se esperaba con sujetos entrenados, para los ejercicios de press de banca y prensa de piernas. Ostrowski y cols. (26) concluyeron que todo lo que se requiere para producir resultados positivos es alcanzar un mínimo de volumen y frecuencia de entrenamiento.

Interesantemente, los resultados de Hass y cols. (25) y de Ostrowski y cols. (26) fueron similares aun con diferentes modalidades de ejercicio, volumen y frecuencia de entrenamiento. A pesar de las grandes diferencias en los protocolos, todos proveyeron un estimulo adecuado y luego un tiempo apropiado para la adaptación. Las rutinas de una serie por ejercicio utilizadas por Hass y cols. (25) y por Ostrowski y cols. (26) demostraron que los sujetos con experiencia en el entrenamiento de la fuerza pueden beneficiarse del entrenamiento de la fuerza con rutinas simples y eficientes en cuanto al tiempo.

La dificultad para extraer conclusiones definitivas acerca de la realización de series múltiples versus series únicas con sujetos experimentados, particularmente dado el pequeño número de estudios, se puede ilustrar adicionalmente por medio de los siguientes análisis que he efectuado utilizando los datos presentados por Wolfe y cols. (6). Los tamaños del efecto de cada uno de los cuatro estudios (23-26) representaron el tamaño del efecto medio a través de los diferentes ejercicios realizados tanto con series únicas como con series múltiples, en sujetos experimentados. Cada estudio proveyó, dos tamaños del efecto, uno para la realización de series únicas y otro para la realización de series múltiples, lo cual es lo más apropiado (9) y no múltiples tamaño del efecto no independientes para cada uno de los seis estudios (21-26), como en los meta-análisis realizados por Rhea y cols. (5) y Wolfe y cols. (6). En el estudio de Ostrowski y cols. (26) he utilizado la media del tamaño del efecto para una y cuatro series ya que la media para el tamaño del efecto para dos y cuatro series era casi la misma. He representado los datos utilizando intervalos de confianza. Dados los datos de los estudios, los intervalos de confianza indicaron el rango dentro del cual caería la media de la población. He utilizado el intervalo de confianza del 95% para indicar que hay un 95% de probabilidad de que la media se encuentra dentro de los rangos señalados. El tamaño del efecto medio y el rango de confianza para estos cuatro estudios (23-26) fueron; utilizando series únicas, 0.44±0.22 (p=0.05) y utilizando series múltiples 0.64±0.25 (p=0.05). Expresados en rangos de confianza (p=0.05), estos cuatro estudios sugieren un rango para el tamaño del efecto de 0.22 a 0.66 para la utilización de series únicas y de 0.39 a 0.89 para la utilización de series múltiples. Estas distribuciones se superponen. Cuando se adicionaron al grupo de estudios los tamaño del efecto medios de los resultados de los tres experimentos realizados por Kraemer (21) utilizando series únicas (0.14) y utilizando series múltiples (0.45), el tamaño del efecto total para la utilización de series únicas fue entonces de 0.38±0.27 (p=0.05) con un rango de 0.11 a 0.65, y para la utilización de series múltiples fue de 0.60±0.26 (p=0.05) con un rango de 0.34 a 0.86, lo cual nuevamente sugiere que las distribuciones se superponen. Cuando se adicionaron las medias para el tamaño del efecto para la utilización de series únicas (0.28) y para la utilización de series múltiples (1.27) del estudio (22) en donde el nivel de entrenamiento de los participantes fue señalado incorrectamente como experimentados, la media total del tamaño del efecto para la utilización de series únicas fue de 0.36±0.26 (p=0.05) con un rango de 0.10 a 0.62, y para la utilización de series múltiples fue 0.71±0.46 (p=0.05) con un rango de 0.25 a 1.17, observándose que las distribuciones aun se superponen. De lo anterior surgen las cuestiones acerca de cuales fueron los estudios incluidos en el meta-análisis, si se extrajeron de cada estudio múltiples tamaños del efecto no independientes o si se extrajo la media del tamaño del efecto de cada estudio, y acerca de como los análisis realizados impactan a las conclusiones extraídas.

Los resultados de Wolfe y cols. (6) son cuestionables debido solamente al pequeño número de estudios que involucraban sujetos con experiencia en el entrenamiento de la fuerza (21-26), Wolfe y cols. no examinaron críticamente cada estudio y además utilizaron múltiples tamaño del efecto no independientes.

Rhea y cols. (7)

Con muy pocos datos reportados, Rhea y cols. (7) afirmaron que su meta-análisis demostraba que el entrenamiento de la fuerza utilizando 3 series era superior a la utilización de 1 serie en sujetos previamente desentrenados y en sujetos entrenados. Estos autores no definieron como se realizó la codificación de los sujetos entrenados y desentrenados. Los desvíos estándar fueron grandes (i.e., ET total=0.28; DE=0.56), indicando una gran variación de los efectos de entrenamiento tanto con una como con tres series. Rhea y cols. (7) incluyeron 16 estudios en su análisis; 12 de ellos fueron citados en una crítica previa (19), y nueve de estos 12 estudios no reportaron diferencias significativas en las ganancias de fuerza como resultado de series únicas o series múltiples (Tabla 4, pag. 17). Uno de los otros tres estudios es el trabajo de Kraemer (21) el cual se demostró previamente que es un estudio con resultados atípicos. De las otras cuatro referencias utilizadas por Rhea y cols. (7), una (27) fue publicada luego de la Declaración de Posición del ACSM (3) y reportó que las ganancias de fuerza en el ejercicio de prensa de piernas eran significativamente mayores en el grupo que realizó 3 series en comparación con el grupo que realizó 1 serie, sin diferencias significativas en la fuerza en press de banca, masa magra corporal, porcentaje de grasa corporal, o en los perímetros del pecho y del muslo. Las otras tres referencias (28-30) fueron

resúmenes en los que ninguno reportaba diferencias significativas en las ganancias de fuerza como resultado del entrenamiento con una o con tres series. Por lo tanto, 12 de los 16 estudios en el meta-análisis (7) no reportaron diferencias significativas entre series únicas y series múltiples. Estos puntos por si solos cuestionan éste meta-análisis que afirma que el entrenamiento con tres series es superior al entrenamiento con una serie por ejercicio.

Rhea y cols. (7) generaron un tamaño del efecto estimado considerando que el grupo que realizó una serie por ejercicio como "grupo control". Los resultados de éste meta-análisis, como ya se señalo, sugieren la superioridad de la utilización de tres series por ejercicio en comparación con la utilización de una serie por ejercicio, con una media para el tamaño del efecto de 0.28±0.56. Rhea y cols. (7) también compararon las medias de los tamaños del efecto que habían igualado la intensidad y controlado la variación (0.70±0.92) con aquellos que no habían igualado la intensidad y la variación (0.20±0.70), y hallaron que esta diferencia se acercaba a la significancia (p =0.12). Los grandes desvíos estándar y un test de significación para la homogeneidad indicaron que otras variables impactaron en los resultados y la probable presencia de estudios con resultados atípicos. Existen ciertas cuestiones acerca de la interpretación de datos en presencia de un test de significación para la homogeneidad, ya que esto significa que hay una gran variabilidad asociada con los efectos del tratamiento (9).

Uno de los principales problemas señalados en la revisión previa del meta-análisis de Rhea y cols. (5) aparece aquí también. En este segundo meta-análisis realizado por Rhea y cols. (7) se generaron un total de 93 tamaños del efecto a partir de 16 estudios. Debido a este problema en su meta análisis, Rhea y cols. (7) no pueden respaldar claramente la superioridad de la utilización de series múltiples.

Peterson y cols. (8)

Peterson y cols. (8) reportaron 37 estudios de entrenamiento que contaban con mediciones de la fuerza pre y post-entrenamiento en atletas competitivos. Estos estudios fueron aparentemente un subconjunto de los estudios reportados en el meta-análisis de Rhea y cols. (5), y los problemas metodológicos señalados en el meta-análisis de Rhea y cols. (5) también aparecen en el meta-análisis de Peterson y cols. (8). Los tamaños del efecto fueron generados para cada resultado de la fuerza, utilizando la siguiente fórmula media post-entrenamiento menos media pre-entrenamiento dividido por el desvío estándar pre-entrenamiento.

Los estudios podrían tener múltiples tamaños del efecto debido a que tienen múltiples resultados tales como ganancias en fuerza en sentadillas, press de banca y dorsal en polea. Como se señaló anteriormente, la utilización de muchos tamaños del efecto no independientes no es apropiada (9).

Peterson y cols. (8) categorizaron el protocolo de entrenamiento en cada estudio en número de series por grupo muscular, entrenamiento a un porcentaje específico de 1RM, y entrenamiento de los grupos musculares dos o tres días por semana. Para cada una de estas variables (series, porcentaje de 1RM, frecuencia de entrenamiento), se extrajeron tamaños del efecto medios de cada uno de los diferentes estudios. Por ejemplo, la utilización de cuatro series por grupo muscular, produjo una media para el tamaño del efecto de 0.90 (±1.23, indicando un muy buen impacto pero un gran desvío estándar, y por lo tanto mucha variación de los resultados) derivado de 119 tamaños del efecto de diferentes estudios. La media para el tamaño del efecto con la utilización de cinco series fue de 0.64±0.73, derivada de 37 tamaños del efecto, y para la utilización de seis series fue de 0.68±0.74, derivada de 26 tamaños del efecto. La media para el tamaño del efecto con la utilización de 8 series fue de 1.22±0.56, pero este resultado estuvo solamente en seis tamaños del efecto, posiblemente a partir de un mismo estudio. Los tamaños del efecto para 12 (0.69), 14 (1.06) y 16 (0.41) series fueron menores pero sin un patrón consistente.

Peterson y cols. (8) afirmaron que su meta-análisis mostró un continuum de incremento de la fuerza para un continuum de incremento en el volumen de entrenamiento, pero los datos reales de su meta-análisis no respaldan esta afirmación. Peterson y cols. (8) también afirmaron que su meta-análisis mostró que los atletas competitivos deberían utilizar ocho series por grupo muscular. Un patrón inconsistente de resultados y el pequeño número de tamaños del efecto para la utilización de ocho series hace que esta conclusión sea cuestionable.

Un patrón inexplicable emerge del análisis realizado por Peterson y cols. (8) en cuanto al entrenamiento con un porcentaje específico de 1RM. El tamaño del efecto para el entrenamiento al 70% de 1RM fue de 0.07±0.06, indicando que no hay impacto sobre la fuerza, mientras que para el entrenamiento al 75% de 1RM, la media del tamaño del efecto fue de 0.73±0.87, para el 80% de 1RM fue de 0.57±0.69 y para el 85% de 1RM fue de 1.12±1.35, los cuales son grandes tamaños del efecto, pero con muy grandes desvíos estándar. Las diferencias tanto en los tamaños del efecto, como en sus desvíos estándar en el rango de entrenamiento de 70-85% de 1RM son confusas. No hay teoría ni mecanismo fisiológico que explique por que un cambio del 5% en la sobrecarga y la realización de unas pocas repeticiones más o unas pocas repeticiones menos pueden resultar en tan grandes diferencias en los resultados.

Este resultado también esta sujeto a cuestionamiento ya que, como se señaló previamente, puede haber una gran variación

interindividual en lo que respecta a cuantas repeticiones hasta el agotamiento puede realizar un sujeto en un ejercicio dado con una RM específica (20).

Peterson y cols. (8) también indicaron que los resultados globales fueron moderados por la utilización de creatina, de entrenamiento periodizado, y de entrenamientos hasta el fallo. Estas variables fueron reportadas en la narrativa como variables que produjeron ganancias de fuerza estadísticamente significativas, pero no se presentaron datos específicos.

Otra limitación fue la forma en que los datos fueron analizados. Peterson y cols. (8) presentaron los resultados en forma de gráficos sin ningún otro análisis formal. A menos que se halla realizado un análisis por separado entre, por ejemplo, los distintos números de series utilizadas, no se puede saber en donde, si es que hubo alguna, están las diferencias significativas. En ausencia de algún patrón obvio, no es posible realizar algún tipo de conclusión. Además, dado que hubo solamente seis tamaños del efecto que contribuyeron a la media para la utilización de ocho series, siendo todos estos tamaños del efecto probablemente de un solo estudio, cualquier conclusión acerca del impacto sobre la fuerza de la realización de ocho series en comparación con cualquier otro tipo de protocolo es altamente cuestionable. Además, la media para la utilización de cuatro series se calculó a partir de 119 tamaños del efecto. Por lo tanto, las dos medias de los tamaño del efecto, cuatro series versus ocho series, no pueden ser comparadas legítimamente.

Sin embargo, para tener una mayor claridad, utilizando los datos del meta-análisis realizado por Peterson y cols. (8), he comparado estadísticamente la diferencia entre cuatro y ocho series por grupo muscular. La diferencia no es significativa ($t<1$).

Los resultados del meta-análisis realizado por Peterson y cols. (8) son altamente cuestionables debido a la utilización de múltiples tamaños del efecto no independientes, a los pocos tamaños del efecto disponibles para cada categoría que afirmaba producir resultados superiores (8 series), el enfoque en la media de los tamaño del efecto y no en la gran variación, y la falta de alguna teoría o mecanismo fisiológico que explicara el patrón de los resultados.

CONCLUSIONES

Cuatro meta-análisis recientes afirmaron que sus resultados mostraban la superioridad de la utilización de series múltiples o la utilización de cuatro o más series por grupo muscular en comparación con la utilización de series únicas por ejercicio en sujetos experimentados o en atletas competitivos. El patrón de los resultados, los tratamientos estadísticos presentados en los estudios, los análisis estadísticos que yo he realizado para tres de los meta-análisis (Rhea y cols. [5]; Wolfe y cols. [6]; Peterson y cols. [8]), la examinación de los estudios incluidos en uno de los meta-análisis (Wolfe y cols. [6]), y la ausencia de cualquier teoría o mecanismo fisiológico, todas estas variables en conjunto en conjunto fallaron en respaldar los protocolos o recomendaciones de entrenamiento de la fuerza con un alto volumen.

La Tabla 1 muestra las normas recomendadas para la realización de un meta-análisis y si cada meta-análisis siguió o no una norma específica. Se puede observar con facilidad que la mayoría de las normas básicas (9) y de las precauciones y advertencias (10-18) para la realización de un meta-análisis y la presentación de los datos no fueron seguidas en los meta-análisis previamente mencionados (5-8).

Tabla 1. *Adherencia a las normas específicas de un meta análisis. ? = no esta claro; * prueba de homogeneidad significativa, pero los resultados no se presentaron.*

Las preguntas principales son “como” y “porque” estos meta-análisis (5-8) fueron publicados con dichos errores. Una respuesta a la pregunta de “como” es que a través de las diferentes disciplinas, los investigadores se han enamorado del meta-análisis a un grado tal que parecen ser capaces de reducir la complejidad de un campo de estudio a un simple número (9, 18). Esta crítica a estos cuatro meta-análisis muestra que cada meta-análisis por si mismo puede ser complejo y que cuando no se siguen las normas (9-18) para la realización de un meta-análisis, los resultados pueden ser incorrectos y falsos.

Una respuesta a la pregunta de “porque” estos estudios fueron publicados es que sus afirmaciones respaldan una faceta

del actual paradigma delineado en la Declaración de Posición del ACSM (3), el cual es que un mayor volumen de ejercicio provoca un mayor incremento en la fuerza. Sin embargo, los datos actuales y algunos estudios dentro de los meta-análisis indican lo opuesto. Contrariamente a la creencia que actualmente prevalece, y que aun no está respaldada, acerca del mayor volumen de entrenamiento (3), un bajo volumen de entrenamiento, protocolos eficientes en cuanto al tiempo, tales como los de series únicas en nueve ejercicios tres veces por semana (25) parecen ser efectivos sin tener en cuenta el nivel de entrenamiento ni los objetivos.

Agradecimientos

Agradezco la retroalimentación y la ayuda con la versión preliminar de este artículo a John Donovan Ph.D, Kevin Davy Ph.D, Jack Finney PhD y a Mike Houston Ph.D. Este artículo estuvo respaldado por una subvención ASPIRES de Virginia Tech.

Dirección para el envío de correspondencia

R.A. Winett, Ph.D., CRHB, Department of Psychology, Virginia Tech, Blacksburg, VA, USA, 24061-0436. (540-231-8747); correo electrónico: rswinett@vt.edu

REFERENCIAS

1. Carpinelli RN, Otto RM (1998). Strength training: Single versus multiple sets. *Sports Med; 26: 73-84*
2. Carpinelli RN (2002). Berger in retrospect: effect of varied weight training programs on strength. *Brit J Sports Med; 36: 319-24*
3. American College of Sports Medicine (2002). Position Stand: Progression models in resistance training for healthy adults. *Med Sci Sports Exerc; 34: 364-80*
4. Lipsey MW (1990). Design Sensitivity Newbury Park. *Sage*
5. Rhea M, Alvar B, Burkett L, Ball S (2003). A meta-analysis to determine the dose-response relationship for strength. *Med Sci Sports Exerc; 35: 456-64*
6. Wolfe BL, LeMura LM, Cole PJ (2004). Quantitative analysis of single- vs. multiple-set programs in resistance training. *J Strength Cond Res 2004; 18: 35-47*
7. Rhea MR, Alvar BA, Burkett LN (2002). Single versus multiple sets for strength: A meta-analysis to address the controversy. *Res Quart Exer Sport; 73: 485-88*
8. Peterson MD, Rhea MR, Alvar BA (2004). Maximizing strength development in athletes: A meta-analysis to determine the dose-response relationship. *J Strength Cond Res; 18: 377-82*
9. Lipsey MW, Wilson DB (2001). Practical Meta-Analysis Thousand Oaks, CA. *Sage*
10. Klein DF (2000). Flawed meta-analyses comparing psychotherapy with pharmacotherapy. *Am J Psychiatry; 157: 1204-11*
11. Lau J, Ioannidis LJ, Schmid CH (1997). Quantitative synthesis in systematic reviews. *Ann Intern Med; 127: 820-26*
12. Shapiro S. (1994). Meta-analysis/Shmeta-analysis. *Am J Epidemiol; 9: 771-78*
13. Sterne JA, Schulz KF, Altman DG, Bartlett C., Egger M (2002). Statistical methods for assessing the influence of study characteristics on treatment effects in meta-epidemiological research. *Stat Med; 21: 1513-24*
14. Glasziou PP, Sanders SL (2002). Investigating causes of heterogeneity in systematic reviews. *Stat Med; 21: 1503-11*
15. Victor N (1995). The challenge of meta-analysis: Discussion. Indications and contra-indications for meta-analysis. *J Clin Epidemiol; 48: 5-8*
16. Matt GE, Navarro AM (1997). What meta-analyses have and have not taught us about psychotherapy effects: a review and future directions. *Clin Psychol Rev; 17: 1-32*
17. Flather MD, Farkouh ME, Pogue JM, Yusuf S (1997). Strengths and limitations of meta-analysis: larger studies may be more reliable. *Control Clin Trials; 18: 568-79*
18. Spitzer WO (1995). The challenge of meta-analysis. *J Clin Epidemiol; 48: 1-4*
19. Carpinelli RN, Otto RM, Winett RA (2004). A critical analysis of the ACSM Position Stand on Resistance Training: Insufficient evidence to support recommended training protocols. *JEPOnline; 7: 1-60*
20. Chromiak JA, Smedley B, Carpenter W, Brown R, Koh YS, Lamberth JG Joe LA, Abadie BR, Altorfer G (2004). Effect of a 10-week strength training program and recovery drink on body composition, muscular strength and endurance, and anaerobic power and capacity. *Nutrition; 20: 420-27*
21. Kraemer WJ (1997). A series of studies the physiological basis for strength training in American football: fact over philosophy. *J Strength Cond Res; 11: 131-42*
22. Kraemer WJ, Ratamess N, Fry AC Triplett-McBride T, Korzis LP, Bauer JA, Lynch JM, Fleck SJ (2000). Influence of resistance training volume and periodization on physiological and performance adaptations in collegiate women tennis players. *Am J Sports Med; 28: 626-33*
23. Kramer JB, Stone MS, O Bryant HS, Conley MS,Johnnson RL, Neiman DC, Honneycutt DR, Hoke TP (1997). Effect of single vs. multiple sets of weight training: impact of volume, intensity, and variation. *J Strength Cond Res; 11: 143-47*
24. Schlumberger A, Stec J, Schmidtbleicher D (2001). Single- vs. multiple-set strength training in women. *J Strength Cond Res; 15:*

284-89
25. Hass CJ Garzarella L, De Hoyos D, Pollock ML (2000). Single versus multiple sets in long-term recreational weightlifters. *Med Sci Sports Exerc; 32: 235-42*
26. Ostrowski KJ, Wilson GJ, Weatherby R, Murphy PW, Little AD (1997). The effect of weight training volume on hormonal output and muscular size and function. *J Strength Cond Res; 11: 148-54*
27. Rhea MR, Alvar BA, Ball SD, Burkett LN (2002). Three sets of weight training superior to 1 set with equal intensity for eliciting strength. *J Strength Cond Res; 16: 525-9*
28. DeHoyos DV, Herring D, Garzarella L, Werber G, Brechue WF, Pollock ML (1997). Effect of strength training volume on the development of strength and power in adolescent tennis players. *Med Sci Sports Exerc; 29 (suppl): S164*
29. DeHoyos D, Abe T, Garzarella L, Hass C, Nordman M, Pollock M (1998). Effects of 6 months of high- or low-volume resistance training on muscular strength and endurance. *Med Sci Sports Exerc; 30 (Suppl): S165*
30. Hass CJ, Garzarella L, DeHoyos DV, Pollock ML (1998). Effects of training volume on strength and endurance in experienced resistance trained adults. *Med Sci Sports Exerc; 30 (Suppl): S115*

Cita Original

Richard A. Winett Meta-Analyses Do Not Support Performance Of Multiple Sets Or High Volume Resistance Training. JEPonline; 7 (5): 10-20, 2004