



Junta de Andalucía
Consejería de Educación y Deporte

Revista Andaluza de Medicina del Deporte

<https://ws072.juntadeandalucia.es/ojs>



Artículo Especial



PRESENT 2020: Texto que desarrolla la lista de verificación para el adecuado informe de la evidencia en ensayos clínicos de deporte y nutrición del ejercicio (Traducción Inglés-Español)

J. A. Betts^a, J. T. Gonzalez^a, L. M. Burke^b, G. L. Close^c, I. Garthe^d, L. J. James^e, A. E. Jeukendrup^e, J. P. Morton^c, D. C. Nieman^f, P. Peeling^g, S. M. Phillips^h, T. Stellingwerff^{i,j}, L. J. C. van Loon^k, C. Williams^e, K. Woolf^l, R. J. Maughan^{ll}, G. Atkinson^m

^a Department for Health, University of Bath, UK

^b Australian Institute of Sport, Belconnen ACT 2616, Australia.

^c Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, UK

^d The Norwegian Olympic and Paralympic Committee and Confederation of Sport.

^e School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, UK

^f Human Performance Laboratory, Appalachian State University, North Carolina Research Campus, Kannapolis, NC 28081, USA

^g School of Human Sciences (Exercise and Sport Science), The University of Western Australia

^h Department of Kinesiology McMaster University Hamilton Ontario Canada

ⁱ Canadian Sport Institute-Pacific, Victoria, BC, Canada

^j Department of Exercise Science, Physical and Health Education, University of Victoria, Victoria, BC, Canada

^k Department of Human Biology, NUTRIM School of Nutrition and Translational Research in Metabolism, Maastricht University Medical Centre+, Maastricht, Netherlands

^l Department of Nutrition and Food Studies, NYU Steinhardt, New York, New York, USA

^{ll} School of Medicine, St Andrews University, St Andrews, UK

^m School of Health and Social Care, Teesside University, Middlesbrough, UK

De: International Journal of Sport Nutrition and Exercise Metabolism, 2020, 30(1): 2-13, <https://doi.org/10.1123/ijsnem.2019-0326>. © Human Kinetics, Inc. Traducido a la lengua española con permiso de los autores y editores, Human Kinetics, Inc. (www.humankinetics.com), que no acepta responsabilidad alguna de la exactitud de la traducción.

Traducción y Adaptación al Español

M. E. Da Silva-Grigolettoⁿ, F. Boscolo Del Vecchio^o, J.D. Beas-Jimenez^p

ⁿ Functional Training Group. Federal University of Sergipe. São Cristovão. Brazil

^o School of Physical Education. Federal University of Pelotas. Pelotas. Brazil

^p Centro Andaluz de Medicina del Deporte. Sevilla. España

INFORMACIÓN DEL ARTÍCULO: Recibido el 13 de noviembre de 2020, aceptado el 13 de noviembre de 2020, *online* el 16 de noviembre de 2020.

Palabras Clave: Rendimiento deportivo; Fisiología ejercicio; Fisiología deporte.

* Autor para correspondencia.

Correo electrónico: J.Betts@bath.ac.uk (J. A. Betts).

<https://doi.org/10.33155/j.ramd.2020.11.004>

Consejería de Educación y Deporte de la Junta de Andalucía. Este es un artículo Open Access bajo la licencia CC BY-NC-ND

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

La guía CONSORT (CONsolidated Standards Of Reporting Trials) de 2010 (<http://www.consort-statement.org/consort-2010>) fue desarrollada para mejorar el informe de ensayos clínicos controlados aleatorios de grupos paralelos, por lo que el cumplimiento de los estándares establecidos se puede demostrar mediante la cumplimentación de la lista de chequeo CONSORT 2010¹. Las principales revistas médicas han respaldado esta iniciativa, que sin duda ha mejorado la realización y el informe de resultados de la investigación clínica y sanitaria. La investigación en el campo de la nutrición deportiva y el metabolismo del ejercicio se beneficiará de estándares similares, pero comúnmente utiliza diseños de investigación diferentes a los ensayos clínicos de grupos paralelos, como los experimentos cruzados.

Se ha publicado recientemente una ampliación de CONSORT que aborda ensayos cruzados aleatorios², con una lista de verificación revisada que se centra en cuestiones de primera relevancia para ensayos clínicos que abordan tratamientos médicos o de asistencia médica. Sin embargo, estas cuestiones específicas de los ensayos clínicos, pueden tener una relevancia diferente, cuando se consideran en relación con los experimentos mecanicistas, estrictamente controlados, realizados en laboratorio, que son habituales en la ciencia del ejercicio. Por ejemplo, los diseños cruzados pueden implicar efectos de orden entre evaluaciones; en los ensayos clínicos, esto suele ocurrir con mayor frecuencia debido al tratamiento o la intervención en sí, lo que requiere un intervalo de lavado suficiente, antes de repetir las evaluaciones. Por el contrario, la tendencia en la investigación sobre nutrición deportiva, se relaciona comúnmente con la evaluación en sí misma, que a menudo, tiende a ser más invasiva o exigente para el participante, que una instantánea sobre su estado de salud. De hecho, las pruebas físicas de rendimiento humano son particularmente propensas a efectos de aprendizaje o fatiga e incluso algunas adaptaciones físicas pueden persistir durante semanas o meses después del primer test. Por ejemplo, el denominado "efecto de intentos repetidos", que describe cómo una sola exposición a un ejercicio físico desacostumbrado que produce daño muscular puede inducir una profunda y duradera protección ante ejercicios similares en el futuro^{3,4}. Los participantes en ensayos de ejercicio también pueden ser atletas de élite, cuyos niveles habituales de actividad física (y dieta) pueden mostrar variaciones profundas a lo largo del tiempo (es decir, periodización), lo que complica aún más la interpretación de estudios longitudinales.

La lista de chequeo PRESENT (*Proper Reporting of Evidence in Sport & Exercise Nutrition Trials*) 2020 (Ver Anexo), por tanto, se ha adaptado de la guía CONSORT para abordar específicamente la combinación única de desafíos y oportunidades que enfrentan los investigadores dentro de los amplios campos de la nutrición deportiva y el metabolismo del ejercicio. El presente documento complementa y desarrolla la lista de verificación CONSORT al poner énfasis y proveer ejemplos comunes o de mayor relevancia para la investigación en esta área temática. La lista de comprobación PRESENT 2020 se diseñó teniendo en cuenta la necesidad de minimizar a los autores la sobrecarga que supone asegurarse de que el informe de sus resultados cumpla con los requisitos de calidad; debería permitir a los investigadores determinar rápidamente si toda la información relevante está incluida en su manuscrito. Por supuesto, es posible cumplir con todos los factores en la lista de verificación a pesar de haber realizado un mal estudio y/o haber informado mal de un buen estudio, mientras que algunos elementos de la lista de chequeo pueden no ser aplicables incluso para investigaciones rigurosamente realizadas. No obstante, la consideración y discusión de los factores identificados en la lista de chequeo debería mejorar la comunicación de los resultados de la investigación relacionada con el ejercicio y la nutrición, en el futuro inmediato, y tiene el potencial de mejorar el diseño y la realización de ensayos a largo plazo.

Las siguientes secciones amplían y justifican cada uno de los elementos incluidos en la siguiente lista de chequeo para envío:

Título

1a El título debe reflejar con precisión los hallazgos principales del estudio, preferiblemente a través de una declaración informativa (por ejemplo, "La cafeína mejora el tiempo de natación de 200 m en nadadores de élite" en lugar de "Los efectos de la ingestión de cafeína en la natación"). La correlación no infiere causalidad, por lo que el lenguaje causal debe reservarse solo para el título de investigaciones experimentales (p. ej., términos como "deteriorado", "producido" o "mejorado"), mientras que la investigación observacional debe emplear declaraciones apropiadas no casuales (p. ej., términos como "relacionado", "correlacionado" o "asociado"). Los títulos deben identificar específicamente la o las variables estudiadas en lugar de medidas relacionadas o indirectas (p. ej., "producción de lactato sanguíneo" o "aclaramiento de glucosa en sangre" no deben establecerse cuando solo se han evaluado las concentraciones sanguíneas de los metabolitos respectivos).

1b El título debe identificar a la población del estudio si las características son directamente relevantes para el diseño del estudio (p. ej., sexo y/o nivel de entrenamiento). Si se utilizan modelos no humanos, se debe indicar la especie.

Resumen

Es posible que algunos lectores no tengan acceso al documento completo, por lo que es imprescindible un resumen bien estructurado y adecuadamente redactado. Los autores deben dar prioridad a la información sobre el estudio actual en lugar de utilizar el resumen para una amplia introducción o justificación.

2a Métodos: se debe resumir la información clave sobre el diseño del estudio, los procedimientos y la población para permitir una comprensión completa del estudio a partir del resumen.

2b Resultados: los lectores están interesados en extraer datos clave que reflejen los principales hallazgos del estudio. El resumen debe presentar datos (por ejemplo, la magnitud absoluta de los valores y el tamaño/precisión de los efectos, especificando qué medidas de tendencia central y variabilidad se establecen) en lugar de simplemente indicar la presencia, ausencia o dirección de los efectos. La presentación de valores de p o estadísticas inferenciales similares no sustituye la presentación de datos reales⁵.

2c Conclusión: Se debe dar prioridad al informe de resultados según la sección anterior, con solo una breve conclusión final a continuación. Se prefiere una conclusión concisa, basada en lo que realmente se midió en el estudio, a interpretaciones especulativas, con un uso prudente del lenguaje para evitar la hipérbole o la inferencia incorrecta de la causalidad⁶. No es apropiado o necesario identificar más prioridades de investigación aquí.

Introducción

3a Esta sección no necesita proporcionar una revisión exhaustiva del área temática, ya que se puede hacer referencia a los artículos de revisión relevantes al presentar brevemente la teoría, los procesos y/o los mecanismos subyacentes. Se debe reconocer la prioridad de publicación: no se deben citar las publicaciones recientes que simplemente confirman los hallazgos anteriores. Por lo tanto, el enfoque principal debe estar en lo que ya se ha hecho, particularmente en los estudios previos más relevantes, asegurando un equilibrio justo de las diferentes

perspectivas siempre que la evidencia actual sea equívoca. Debe quedar claro para el lector qué información nueva pretende proporcionar el estudio y por qué esa información es importante. Lo primero significa más comúnmente identificar la novedad (es decir, qué mostrará este estudio que no se haya mostrado antes) pero también podría significar justificar por qué es necesaria la replicación de un estudio anterior. Si el elemento novedoso de un estudio se basa en una larga lista de calificadores (por ejemplo, el efecto de A sobre B ya se ha establecido, pero no dentro de la población C y el contexto D), entonces la Introducción debe justificar por qué esos calificadores son interesantes y merecen una investigación más a fondo. (es decir, ¿qué es lo importante de la población C y/o el contexto D?). Por lo tanto, el lector debe tener una comprensión clara de la justificación del trabajo.

3b Una hipótesis formal puede no ser siempre apropiada para trabajos de naturaleza exploratoria o cualitativa, pero la Introducción debe concluir con alguna forma de objetivo claro o pregunta de investigación (es decir, suficiente para expresar la pregunta central y presentar las principales variables bajo investigación).

Método

4 Ética

Los detalles con respecto a la aprobación ética deben proporcionar una referencia adecuada al organismo de aprobación para permitir la verificación de la aprobación si es necesario. Se anima a los autores a proporcionar el número de aprobación del comité de ética (si está disponible) y la fecha de aprobación. El consentimiento informado debe registrarse explícitamente para los ensayos en humanos cuando sea relevante, y justificarse cuidadosamente para cualquier ensayo en el que se justifique el engaño o las pruebas de los participantes sin consentimiento previo⁷. En el caso de participantes menores de la edad legal de consentimiento, se debe obtener su consentimiento y el consentimiento de sus padres o tutores legales. Se recomienda encarecidamente el registro, previo a la prueba, de los estudios que se someten a evaluación como "Ensayos Clínicos", y se deben proporcionar detalles sobre el acceso al Registro de Ensayos Clínicos. Cuando un estudio aporta datos recopilados durante el seguimiento de rutina de personas por su condición empleado o participante (por ejemplo, atletas o poblaciones militares), una declaración clara sobre la fuente de datos y las razones de la falta de aprobación previa por parte de un comité, como su requisito como empleado, debe hacerse explícitamente⁸. En tales situaciones, puede ser apropiado introducir cierto grado de cegado del proceso, tales como una clave de registro para mantener el anonimato y/o la confidencialidad de los datos.

5 Diseño

5a Los detalles completos del diseño de la investigación deben resumirse al principio de la sección Métodos para establecer el contexto para el lector, utilizando terminología y nomenclatura comunes (por ejemplo, ensayo paralelo/cruzado, aleatorizado, contrabalanceado, cegamiento, observacional, etc.), ambos para proporcionar una descripción general de primera línea desde el principio y para facilitar la extracción precisa de datos para aquellos que realizan futuras revisiones sistemáticas y metaanálisis. Todo el trabajo descriptivo, incluidas las encuestas y los estudios de casos, debe identificarse como investigación observacional. Si los datos presentados en el documento provienen de un análisis secundario de un estudio más amplio/anterior en lugar de una investigación primaria, esto también debe identificarse claramente en la información introductoria sobre el diseño del estudio, con comentarios sobre si el estudio fue diseñado de manera específica o adecuada para los

finés informados en el presente documento y si los datos se replican de una publicación anterior.

6 Muestreo

6a Deben explicarse los métodos de reclutamiento, los criterios de elegibilidad (inclusión/exclusión) y los métodos de muestreo, proporcionando preferiblemente una justificación para la población diana. En particular, los autores de estudios de casos deben justificar el interés o la elección de su(s) participante(s) único(s). Si los autores del estudio reclutan intencionalmente una proporción no representativa de hombres y mujeres, esto debería estar justificado.

6b Para proporcionar una transparencia total en torno a la población a la que se pueden generalizar los hallazgos del estudio, es esencial una descripción exhaustiva de las características de los participantes. Todos los datos deben informarse con la precisión adecuada en relación con la herramienta de medición y las unidades en cuestión⁹; por ejemplo, no es necesario registrar o expresar la edad en años decimales o con mayor precisión que el número entero (por ejemplo, es apropiado informar una edad en años como 25 en lugar de 25.24). La elección de qué variables demográficas, antropométricas o de estilo de vida se deben considerar en términos de relevancia para la interpretación y generalización de los ensayos dentro de este campo (por ejemplo, estado de entrenamiento, nivel competitivo, habituación a la dieta). El término "élite" se usa comúnmente cuando se describe a los atletas: los términos descriptivos deben usarse de manera apropiada y calificada por información objetiva (por ejemplo: mejor deportista de una modalidad/evento relevante, nivel de representación internacional/nacional, ranking o clasificación de puntos donde sea relevante, puntuación actual de una batería de test de calidad contrastada, etc.).

6c El escenario del estudio debe describirse tanto en términos de pruebas de laboratorio *versus* pruebas de campo como de la región/sitio (s) de recogida de datos. Si se recopilan datos en múltiples laboratorios independientes, se deben proporcionar detalles de las mediciones que se realizaron en cada ubicación, con datos relevantes de confiabilidad específicos para cada sitio. También se requieren fechas de inicio y finalización de las pruebas. Los datos ambientales apropiados (es decir: temperatura, humedad relativa, altitud) deben establecerse cuando sean relevantes para la interpretación del estudio.

6d Se debe justificar el tamaño de la muestra que se reclutó y analizó. La naturaleza pragmática de los experimentos en este campo, que requieren mucho trabajo/tiempo y a menudo son invasivos, es tal que el tamaño de la muestra, a veces, depende de la disponibilidad de voluntarios y/o recursos humanos. Sin embargo, también debemos considerar que puede ser poco ético incluir más participantes de los necesarios para probar la hipótesis. Por lo tanto, se debe proporcionar una estimación de la potencia/precisión estadística basada en el mínimo efecto necesario y en el tamaño muestral disponible¹⁰ y todos los componentes de los cálculos deben estar claramente justificados e informados con suficiente detalle para permitir la réplica (es decir, alfa, beta, mínimo efecto necesario, estimación de la varianza, por ejemplo, la desviación estándar antes y después del cambio). Deben evitarse las estimaciones de potencia *post-hoc* (basadas en el tamaño del efecto observado en el propio estudio). También debe expresarse claramente si la hipótesis es de naturaleza "superior", donde se supone que un tratamiento/intervención difiere del control u otra intervención, o si el estudio se parece más a un "ensayo de equivalencia" en el que un tratamiento/intervención se considera que tiene una efectividad similar a otra intervención, pero quizás sea, de alguna manera, más simple, menos costosa o menos invasiva. En la última situación, el procedimiento de prueba de hipótesis nula no es apropiado: nunca debe usarse $p > 0.05$ para inferir que "no hay diferencia" entre dos o más estimaciones de muestra¹¹ y tampoco

deben interpretarse valores *p* marginalmente superiores a 0.05 como una "tendencia".

Las estimaciones del tamaño de la muestra pueden basarse en pruebas de hipótesis nulas o en la amplitud deseada del intervalo de confianza, pero la elección debe estar justificada. La sección Discusión puede proporcionar a los autores más oportunidades para enmarcar sus inferencias estadísticas, junto con otras preocupaciones que puedan surgir sobre la replicación de los resultados del estudio. Por ejemplo, la noción de que un tamaño de efecto observado es grande simplemente debido al error de muestreo asociado con una muestra relativamente pequeña, un problema conocido como la "maldición del ganador", puede ser sustentado por estudios estadísticos como el "Valor Sorpresa (S)"¹². Es importante destacar que los tamaños de muestra inadecuados no se justifican necesariamente simplemente porque la población en cuestión es rara o difícil de reclutar (por ejemplo, atletas de élite) o porque se han utilizado tamaños de muestra similares en publicaciones anteriores; tales casos pueden presentarse mejor como trabajo piloto o estudios de casos o, alternativamente, el tamaño de la muestra podría complementarse al incluir datos de una población más amplia (por ejemplo, atletas subélite), a menos que exista una razón para esperar que el resultado principal responda claramente en la muestra más enfocada.

7 Intervenciones

La intervención/tratamiento es posiblemente el elemento más importante del informe de metodología; incluso los detalles aparentemente triviales de la variable independiente pueden ser críticos para una comprensión adecuada de las respuestas atribuidas a los grupos/condiciones. Para compuestos nutricionales con ingredientes complejos, a menudo se recomienda una comparación por pares (posiblemente tabulada) para obtener informes claros y completos. Críticamente, es totalmente inadecuado proporcionar simplemente un nombre de fabricante/producto para un informe científico porque la composición real de dichos productos disponibles comercialmente puede ser patentada, reformulada, discontinuada y/o diferente de la composición informada de cada lote producido; el valor científico prospectivo de los resultados depende enteramente de que el lector pueda establecer en qué consistió la intervención. Por la razón anterior, debe considerarse cuidadosamente si el compuesto nutricional bajo investigación requiere un análisis para verificar que los ingredientes informados están presentes en las cantidades indicadas y que los ingredientes o contaminantes no revelados están ausentes (en algunos casos puede ser que ciertos medidas fisiológicas se pueden usar para verificar la eficacia del protocolo de suplementación y, por lo tanto, pueden verificar la mera presencia o ausencia de ingredientes clave entre tratamientos). Los productos simples que se formulan internamente (por ejemplo, soluciones de glucosa preparadas de acuerdo con un protocolo claro) pueden no requerir tal verificación, pero la investigación transparente y creíble de los productos disponibles comercialmente a menudo requiere tales verificaciones del lote específico utilizado en el proyecto de investigación. Por último, cuando se seleccionan, compuestos o mezclas nutricionales particulares, para su examen, puede ser interesante explicar la justificación de las decisiones con respecto a la elección de la intervención en relación con los resultados (por ejemplo, ver la subsección titulada Approach to the Research question en Bailey et al.¹³).

8 Mediciones

8a Deben especificarse claramente, las medidas de los principales resultados, los destinados a responder las cuestiones planteadas en la investigación. Las eventuales conclusiones e interpretaciones, deberían basarse en los resultados principales

obtenidos (especialmente en el caso de resultados nulos), en lugar de centrarse en aquellos resultados que respondieron mejor o fueron más consistentes con la hipótesis. La categorización de las variables que se miden en una escala continua ("dicotómica") debe evitarse a menos que haya una razón sólida para hacerlo¹⁴.

8b Se debe proporcionar una justificación para las herramientas de medición seleccionadas. Las medidas relacionadas con el análisis de la dieta y/o los tests de ejercicio, son frecuentemente utilizados en la investigación sobre nutrición deportiva. Con respecto a lo primero, el método de evaluación dietética debe describirse con precisión (por ejemplo, autoinforme *versus* medida directa, recogida prospectiva *versus* retrospectiva, ponderación *versus* estimación de la cuantificación mediante medidas caseras) y deben considerarse las limitaciones, junto con detalles completos del nombre y versión de cualquier programa dietético utilizado. Con respecto a esto último, la información de interés incluye la naturaleza de la prueba (por ejemplo, capacidad de ejercicio/tiempo hasta el agotamiento *versus* rendimiento físico/duración de la prueba), la familiaridad/familiarización de los participantes con el protocolo de prueba y si la intensidad es absoluta o relativa a otro parámetro (p. ej. % $\dot{V}O_{2\text{máx}}$). Los protocolos cuyo objetivo es medir el rendimiento físico deben ser válidos, fiables y sensibles¹⁵. En todos los casos, el grado de error de medición, asociado a cada resultado, debe expresarse utilizando estadísticos relevantes de confiabilidad¹⁶. Para los ensayos bioquímicos, los datos de confiabilidad deberían derivarse idealmente de un análisis propio, en lugar de utilizar la confiabilidad intra o entre ensayos, informada por el fabricante. Los investigadores deben tener en cuenta si los analizadores químicos de uso clínico destinados al uso diagnóstico, proporcionan la precisión necesaria para fines de investigación.

8c Los autores deben identificar y justificar claramente la diferencia más pequeña entre los tratamientos que se considera significativa, que debe ser coherente con las estimaciones del tamaño de la muestra que se expone en el ítem 6d y que debe tenerse en cuenta al interpretar los datos (es decir, un efecto menor que el considerado significativo, no debería sobreinterpretarse luego, incluso si resulta ser estadísticamente significativo). Es importante tener en cuenta la distinción entre el efecto o asociación de menor valor y el menor efecto detectable¹⁷. El primer valor, a menudo denominado "diferencia objetivo" de "diferencia mínima clínicamente importante" (MCID) es el valor de cambio o asociación que se considera importante para los participantes, atletas o pacientes. Idealmente, este umbral se basaría en las relaciones conocidas entre el cambio en el resultado de interés y el cambio en el rendimiento deportivo real, la adaptación fisiológica o la morbilidad/mortalidad en el contexto de los ensayos clínicos. Sin embargo, dicho conocimiento puede ser difícil de obtener. Remitimos a los lectores a las publicaciones DELTA y DELTA2 para un tratamiento completo de este tema, incluidos todos los enfoques para llegar a un MCID^{18,19}.

El cambio mínimo detectable es un umbral basado en la probabilidad de que un cambio, en un atleta o paciente concreto, se deba a un error aleatorio dentro de los sujetos (test-retest) o no, con cierta probabilidad de que ocurra. Si bien esto puede ser útil para el seguimiento de las personas, es importante tener en cuenta que el cambio detectable más pequeño y la MCID pueden ser de diferente magnitud. Para fines de investigación, el MCID y los enfoques para obtener este valor son los más importantes. Del mismo modo, con respecto a las correlaciones, la interpretación siempre debe considerar la fuerza, la dirección y la forma de una relación; por ejemplo, una relación lineal débil, inversa (p. ej., $r = -0.1$) puede tener un significado o una utilidad limitados independientemente de si es estadísticamente significativa.

9 Aleatorización

La aleatorización en un ensayo puede ser un componente crítico de la validez experimental, pero es un apartado donde los detalles

importantes a menudo se omiten. Se deben proporcionar detalles completos para identificar, con precisión, cómo y por quién se generó e implementó la secuencia de asignación aleatoria²⁰⁻²². Por ejemplo, dichos detalles podrían incluir el tipo de secuencia de asignación aleatoria, la relación de asignación y los detalles de cualquier minimización o restricción (por ejemplo, estratificación, bloqueo y tamaño de bloque), con una identificación clara de los individuos responsables de: generar/administrar la secuencia de asignación, reclutar/inscribir voluntarios, asignar tratamientos, realizar la recogida de datos y analizar las muestras.

10 Cegamiento

Si bien los experimentos se diseñan e inician comúnmente en base a una hipótesis de trabajo en relación con la respuesta anticipada a una pregunta de investigación preestablecida, se ha reconocido desde hace más de 150 años que las mediciones deben realizarse sin ninguna idea preconcebida o suposiciones a priori con respecto a los resultados esperados²³. Este razonamiento forma la base del concepto moderno de cegado experimental, cuyo objetivo es controlar los sesgos conscientes o subconscientes de los investigadores, participantes y/o evaluadores de los resultados al mantenerlos al margen de la respuesta esperada a un estímulo dado²⁴. El cegamiento es una técnica que permite a los investigadores llevar a cabo el estudio sin expectativas preconcebidas de los resultados, a pesar de la existencia de una hipótesis de trabajo. Este ideal se puede lograr ocultando la asignación del tratamiento y/o el propósito del experimento a cualquier persona involucrada que pueda influir en las mediciones que se realicen.

Aunque el cegamiento es importante a varios niveles, cuando las medidas de los resultados están sujetas a un control consciente o subconsciente, el cegamiento de los participantes se convierte en un tema crítico. Los ejemplos de tales medidas en el campo de la nutrición deportiva y el metabolismo del ejercicio incluyen métricas sobre el rendimiento físico (por ejemplo, contracción muscular voluntaria máxima, tiempo de fatiga en una tarea de ejercicio) o percepciones subjetivas (por ejemplo, calificación de esfuerzo percibido o dolor). Estas medidas de los resultados podrían compararse con variables fisiológicas objetivas, que generalmente son menos propensas a los efectos de expectativa/placebo (por ejemplo, cambios en las concentraciones sistémicas de metabolitos o la utilización del sustrato durante el ejercicio). Sin embargo, algunas intervenciones son difíciles o esencialmente imposibles de cegar verdaderamente (por ejemplo,

crioterapia o ejercicio) y, en algunos casos, la conciencia de los participantes sobre la intervención es fundamental para el mecanismo de acción propuesto (por ejemplo, intervenciones psicológicas). En tales situaciones, donde el cegamiento de la asignación del tratamiento es imposible o indeseable, puede justificarse cegar a los participantes para el propósito del experimento, evaluando la medida primaria de los resultados de manera encubierta, dentro de un contexto de estudio más amplio. Por ejemplo, a los participantes se les puede ofrecer una comida al finalizar las pruebas, después de lo cual se puede buscar el consentimiento *post-hoc* para determinar la ingesta natural de alimentos ad limitum, en función de cualquier alimento sobrante/no consumido. En la [Tabla 1](#) se proporcionan ejemplos de los diversos niveles de cegamiento.

Al realizar un estudio cegado, es especialmente importante evaluar el éxito del cegamiento. Por ejemplo, un tratamiento puede implementarse de manera ciega, por lo que los participantes no son informados de la asignación del tratamiento y se puede hacer todo lo posible para ocultar la asignación del tratamiento utilizando condiciones de placebo/control que coincidan con el sabor y la textura, utilizando contenedores opacos, etiquetas anónimas, tratamientos simulados, o incluso los nuevos métodos de administración (por ejemplo, administración nasogástrica de nutrientes, Funnell et al.²⁵). Sin embargo, el ensayo solo está verdaderamente cegado, en el sentido de que el sesgo consciente/subconsciente se controla, si los participantes desconocen su asignación al tratamiento. Por lo tanto, más allá de los métodos anteriores destinados a ocultar la asignación del tratamiento, el éxito del cegamiento se puede evaluar fácilmente mediante un cuestionario formal de salida. Por ejemplo, para un diseño cruzado en el que cada participante tiene la oportunidad de experimentar y comparar condiciones, esto se puede lograr a través de tres respuestas binarias directas 'sí/no': a) ¿el participante (o experimentador/evaluador) detectó alguna diferencia entre tratamientos? en caso afirmativo: b) ¿se sintieron capaces de identificar su asignación de tratamiento? en caso afirmativo: c) ¿identificaron correctamente su asignación de tratamiento? Para las intervenciones en las que los participantes pueden detectar diferencias entre tratamientos, también puede ser adecuado examinar sus conocimientos previos, preconcepciones y creencias sobre los efectos esperados del tratamiento.

Finalmente, la investigación solo debe informarse como cegada si el investigador confía en que los participantes no estaban al tanto de la asignación del tratamiento; si se deduce que una

Tabla 1. Resumen de varias formas de cegamiento experimental.

| Tipo de Cegamiento | Descripción | Justificación |
|--------------------|--|--|
| Abierto | Todas las categorías de individuos (investigadores, participantes y evaluadores de resultados) saben quién ha recibido qué intervención a lo largo del ensayo. | Algunas intervenciones son imposibles de cegar verdaderamente (por ejemplo, ejercicio) o la conciencia de los participantes sobre la intervención es inherente a la pregunta de investigación |
| Simple Ciego | Una categoría de individuos (normalmente los participantes pero potencialmente los investigadores o evaluadores de los resultados) desconoce quién recibió qué intervención durante el ensayo. | Cuando los participantes están cegados: <ul style="list-style-type: none"> • Es menos probable que se tengan respuestas psicológicas o físicas sesgadas a la intervención. • Es menos probable que se busquen intervenciones complementarias adicionales. • Es más probable que se cumpla con la intervención. • Es menos probable que se abandone la prueba sin proporcionar datos. |
| Doble Ciego | Investigadores, participantes y evaluadores de los resultados desconocen las asignaciones de intervención durante el ensayo. | Cuando los investigadores/evaluadores están cegados: <ul style="list-style-type: none"> • Es menos probable que se transfieran inclinaciones o actitudes a los participantes. • Es menos probable que los sesgos afecten a las evaluaciones de los resultados, especialmente con resultados subjetivos. |
| Triple Ciego | Los investigadores, participantes y evaluadores de los resultados desconocen las tareas de intervención durante el ensayo y durante el análisis de datos. | Es menos probable que se tengan sesgos que afecten los análisis estadísticos y la interpretación. |

proporción de la cohorte general pudo distinguir los tratamientos, entonces puede ser adecuado informar en qué medida las respuestas observadas pueden haber sido influenciadas por el éxito del cegamiento. Si bien el cegamiento puede reducir el sesgo en las investigaciones, no supera la falta de aleatorización, que es un error común²⁴. Además, si el cegamiento completo de un tratamiento es difícil o imposible y el rendimiento del ejercicio es el resultado principal, entonces las pruebas de ejercicio deben seguir las recomendaciones para evitar los comentarios relevantes sobre el rendimiento durante las pruebas de ejercicio¹⁵.

11 Estandarización

La estandarización del comportamiento y el entorno de los participantes antes de las pruebas y durante la captura de datos, puede reducir la variabilidad en las medidas basales tanto entre como intra individuos. Esto puede aumentar la capacidad de detectar efectos verdaderos con tamaños de muestra relativamente pequeños, aumentando efectivamente la relación señal/ruido. Las consideraciones clave incluyen la duración de los controles previos a la prueba y si la estandarización se realiza entre diferentes participantes (es decir, controlando la variación interindividual) o solo dentro de los participantes (es decir, controlando la variación intraindividuo), lo que depende en gran medida, de si la investigación implica un diseño paralelo o cruzado. Los parámetros comunes en torno al control del comportamiento/características de los participantes, dentro del campo de la nutrición deportiva y el metabolismo del ejercicio, incluyen actividad física, dieta/hidratación, uso de medicamentos/suplementos y ciclo menstrual. Debe establecerse claramente si este control fue facilitado directamente por los investigadores (por ejemplo, proporcionar comidas) o, de no ser así, si se emplearon evaluaciones objetivas para verificar la replicación exitosa (por ejemplo, monitorización de la frecuencia cardíaca para confirmar la abstinencia del esfuerzo físico intenso antes de la prueba). Si dicho monitoreo se ha completado, a menudo es relevante informar los datos para proporcionar un contexto futuro al contrastar los resultados con otros ensayos. Por ejemplo, incluso si la dieta coincidiera estrechamente entre las condiciones, los lectores pueden beneficiarse al saber si las intervenciones nutricionales se contrastan con una dieta de fondo que sea relativamente alta o baja en energía u otros nutrientes. Un resumen más detallado de los métodos para implementar e informar las técnicas de estandarización dietética se expone en Jeacocke y Burke²⁶.

La influencia de la presencia y de las etapas del ciclo menstrual debe considerarse cuando sea relevante: existen numerosos enfoques para controlar este factor; cada uno potencialmente justificado en función de un equilibrio de validez interna y externa. Se debe proporcionar una justificación para explicar si la fase menstrual se controló en o entre las participantes y qué fase del ciclo se seleccionó para minimizar las influencias confundentes y/o por ser la más representativa. Si se establece una fase específica del ciclo menstrual, debe quedar claro si esto se basó en fechas autoinformadas o en mediciones de hormonas sexuales.

Para obtener más información sobre los beneficios y las técnicas para la estandarización previa y durante la prueba de factores como el ejercicio reciente, la aclimatación, el ruido/las distracciones, el estímulo y la conciencia del paso del tiempo o los resultados anteriores, se remite al lector a otras revisiones^{15,27}.

12 Efectos del orden

Debe analizarse e informarse del potencial efecto del orden entre los ensayos en serie del mismo participante, particularmente durante los estudios que utilizan un diseño cruzado de medidas repetidas. Críticamente, si una evaluación inicial bajo una condición ejerce efectos duraderos o adaptaciones

que exceden el intervalo de lavado antes de la siguiente evaluación para la siguiente condición, entonces la aleatorización y la compensación del orden de tratamiento no previenen el problema de los efectos de orden/periodo/secuencia. Este escenario es particularmente problemático si existe alguna interacción entre el efecto de la intervención y la magnitud de los efectos de arrastre entre mediciones; de hecho, si los efectos del orden son evidentes, entonces ningún control puede rectificar el problema y aislar los efectos de la intervención. Los ejemplos de medidas preventivas para minimizar la probabilidad de efectos del orden incluyen: permitir un lavado adecuado entre evaluaciones; proporcionar una adecuada familiarización previa con las pruebas y/o muestrear una población que ya esté acostumbrada al protocolo y, por lo tanto, es poco probable que exhiba un marcado aprendizaje o efectos de fatiga entre las exposiciones (por ejemplo, el uso de atletas entrenados para las pruebas de ejercicio). Sin embargo, períodos de lavado más prolongados entre las evaluaciones (es decir, semanas a meses) podrían inducir efectos de entrenamiento/desentrenamiento y/o introducir una variación estacional como factores de confusión.

Los experimentos en los que los efectos del orden son evidentes no son necesariamente inadecuados para su publicación, pero el efecto del orden debe mencionarse claramente al informar los resultados. El análisis de los datos para los efectos del orden puede tomar la forma de aplicar técnicas estadísticas similares a las utilizadas para el análisis primario entre condiciones pero con referencia a las evaluaciones secuenciales²⁸, con la debida consideración de que los efectos de la intervención y del orden de tratamiento pueden interactuar (por ejemplo, el placebo puede ser consistentemente inferior solo cuando se aplica en el primer ensayo). Sin embargo, dadas las complejidades de tales análisis estadísticos y que la ausencia de diferencias detectables entre pruebas secuenciales no necesariamente infiere que fueron equivalentes, puede valer la pena inspeccionar los datos individuales en busca de evidencia de cualquier cambio sistemático de acuerdo con el orden del ensayo.

13 Estadísticas

13a Deben identificarse las comparaciones que informarán las inferencias principales sobre los datos del estudio, con los métodos estadísticos descritos con detalladamente. En particular, debe existir consistencia entre el enfoque propuesto, la justificación presentada en la Introducción y los datos informados. Por ejemplo, si el estudio ha sido diseñado para contrastar las respuestas a lo largo del tiempo a una intervención *versus* un placebo, el análisis debe enfocarse en las diferencias entre esas condiciones en lugar de la presencia o ausencia relativa de cambios desde el inicio dentro de cada condición²⁹.

13b Cualquier análisis no planificado o exploratorio debe distinguirse claramente del propósito y análisis principales, con justificación para cualquier análisis intermedio o condiciones potenciales bajo las cuales un ensayo podría detenerse, incluida la consideración de análisis por intención de tratar *versus* análisis por protocolo. En particular, tenga cuidado con el análisis de subgrupos con estratificación al inicio del estudio porque la regresión a la media puede comprometer las inferencias³⁰. Del mismo modo, estratificar los resultados del estudio para el análisis de respondedores/no respondedores está lleno de dificultades y peligros³¹. Estas y otras dificultades estadísticas frecuentes se enumeran junto con ejemplos relevantes, explicaciones adicionales y posibles soluciones en la [Tabla 2](#).

13c Los investigadores deben conocer los supuestos subyacentes a cada enfoque estadístico seleccionado e informar claramente cómo se manejó cualquier apartado de estos supuestos. Los datos más importantes para verificar la normalidad aproximada son típicamente los residuos del modelo estadístico seleccionado. En el contexto de una prueba t pareada, estos datos son las puntuaciones o diferencias de cambio emparejadas

individualmente. Además, los análisis paramétricos a menudo pueden ser robustos, incluso cuando los datos en bruto no se distribuyen normalmente. Cualquier transformación de los datos debe estar claramente justificada³². En general, la transformación logarítmica de los datos, sin verificar primero si estos lo requieren o se benefician de esta transformación, no es una buena práctica. Los investigadores también deben ser conscientes de que las pruebas estadísticas de asimetría (donde la hipótesis nula es que los datos están normalmente distribuidos) tienen tendencia a dar problemas cuando los tamaños de muestra son pequeños. Similar al diseño de series de tiempo, descrito en el ítem 13a, en el que hay medidas repetidas sobre una serie de puntos en el tiempo para cada condición, también puede ser necesario ajustar las diferencias de la línea de base, en la medida del resultado principal³³. A este respecto, los autores deben interpretar cuidadosamente cualquier diferencia en el cambio previo al post,

cuando existan diferencias significativas entre los grupos o las condiciones en los resultados al inicio del estudio; tal problema puede no corregirse calculando los cambios porcentuales desde el inicio o interpretando un término de interacción grupo/condición x tiempo³⁴.

Resultados

14 Flujo de Participantes

14a Dependiendo de la complejidad y la naturaleza del diseño experimental (por ejemplo, seguimientos múltiples con poca adherencia, muchas pérdidas de seguimiento y exclusiones), puede ser apropiado presentar un diagrama de flujo del número de participantes que fueron reclutados, elegibles, inscritos, aleatorizados, asignados, recibido tratamiento y analizado en cada

Tabla 2. Errores estadísticos comunes en los ensayos de nutrición deportiva y metabolismo del ejercicio.

| Dificultad encontrada | Ejemplo | Explicación | Soluciones |
|---|--|---|---|
| Agrupar datos de múltiples grupos o condiciones y analizarlos como una sola muestra combinada Ver: Bland ^{35,36} | Agrupando datos de 10 participantes en 3 condiciones de dieta y examinando la correlación entre dos variables medidas, como si n = 30. | Esta es una "pseudoreplicación", por la cual el tamaño de la muestra se infla erróneamente y hay confusión entre las inferencias dentro de los sujetos y entre sujetos. | Cualquier factor de agrupación debe estar presente en el modelo estadístico. Las réplicas se pueden promediar dentro de cada participante para cuantificar las correlaciones entre sujetos, mientras se conservan los grados correctos de libertad. Del mismo modo, todos los "grupos" (por ejemplo, sitio de laboratorio) deben incluirse en el modelo estadístico. |
| Comparación de cambios/respuestas no ajustadas (en bruto) entre diferentes ramas o subgrupos del ensayo Ver: Vickers ^{33,34} | Comparar los cambios previos y posteriores entre dos grupos de edad diferentes o hombres/mujeres, porque estos grupos pueden diferir al inicio | La regresión a la media puede influir en las inferencias cuando se forman subgrupos sobre la base del estado inicial u otros factores que pueden conducir a un desequilibrio inicial para la variable medida del estudio. | Un modelo de análisis de covarianza (ANCOVA) es apropiado mediante el cual los valores de referencia para cada grupo se incluyen en el modelo estadístico como una covariable. Se ha informado que ANCOVA es superior a la comparación de cambios porcentuales entre grupos o al cálculo de un término de interacción grupo x tiempo. |
| Formando grupos o "condicionamiento" sobre la base de la variable de resultado y cuantificando las diferencias grupales en los valores basales Ver: Atkinson ^{16,37} | Comparación de los valores de referencia entre los grupos de "respondedores" y "no respondedores" | Esto crea un problema similar de regresión a la media como en el caso anterior, lo que hace que los respondedores sean más bajos al inicio (y viceversa) debido a este artefacto estadístico. | Existen muchas dificultades en el "recuento de respondedores" y en el análisis de factores que pueden influir en la respuesta individual. El condicionamiento a la magnitud de la respuesta rara vez es apropiado. |
| Correlacionando el estado inicial propuesto u otros predictores con la respuesta al tratamiento solo en el grupo de intervención. Ver: Heckdteden ³⁸ y Lonergan ³² | Correlación del VO2máx basal con la línea base para el seguimiento del cambio en el VO2máx en respuesta al entrenamiento | La regresión a la media podría influir aquí nuevamente en las inferencias. | Cualquier predictor de respuesta debe modularse como una interacción predictor x rama del ensayo con el ajuste covariable para los valores basales del resultado (porque el predictor y el resultado pueden estar ya correlacionados al inicio). |
| Correlacionar variables que ya están relacionadas matemáticamente Ver: Bland ⁴⁰ , Tu ⁴¹ | Correlacionando el tiempo total de triatlón con el tiempo de la fase de ciclismo, o correlacionando el consumo de oxígeno (ml · kg-1 · min-1) con el índice de masa corporal (kg · m-2). | Esto es "acoplamiento matemático" o "relacionar una parte con un todo" por el cual una variable ya es un componente en el cálculo de otra variable. Tales correlaciones pueden ser <i>espureas</i> . | Esta práctica debe ser evitada. A veces, las correlaciones parciales y las asociaciones ajustadas multivariadas pueden ayudar a "ajustar" la correlación entre dos variables cuando hay una "tercera variable" que se considera que influye en ambas variables, pero esta tercera variable no debe ser inherente a los cálculos de las otras dos variables. |
| Informar diferencias individuales en la respuesta solo en los datos del grupo de intervención/tratamiento Ver: Atkinson ^{31,37} | Trazar la respuesta del VO2máx al entrenamiento para cada individuo en el grupo de intervención, o informar la desviación estándar de la respuesta solo para el grupo de intervención. | Esta práctica no reconoce que pueden ocurrir en el cambio diferencias individuales similares en el grupo de control en el contexto "contrafactual" en los ensayos controlados aleatorios. | La variación del cambio del grupo de control debe incluirse en el enfoque de análisis para cuantificar cualquier heterogeneidad de respuesta individual hipotética. |
| Llegar a conclusiones únicamente sobre la base de valores p estadísticamente significativos (o no significativos). Ver: Greenland ⁴² | Informar que una mejora en el rendimiento es estadísticamente significativa, pero por un margen que sea irrelevante para los atletas, o viceversa. | Aquí, la relevancia clínica o práctica de la magnitud de los efectos no se considera adecuadamente. | Un efecto mínimo importante debe racionalizarse <i>a priori</i> y compararse con el tamaño del efecto observado. Los intervalos de confianza al 95% deben informarse por rutina. Un valor p no significativo nunca debe usarse para concluir "sin efecto" o "sin diferencia", ya que el proceso de prueba de hipótesis nula no está diseñado para hacer inferencias sobre equivalencia. |

VO2máx: Consumo máximo de oxígeno.

etapa de tiempo. Para diseños más directos, puede ser posible simplemente comunicar esta información dentro del texto, pero sigue siendo esencial para identificar claramente el número de participantes (denominador) incluido en cada análisis (es decir, se requiere para cada etapa de tiempo y variable de resultado específicos). En particular, se deben proporcionar razones claras para cualquier exclusión del conjunto de datos final, indicando y justificando los cambios de la asignación del tratamiento original⁴³.

14b Los investigadores deben evitar agrupar los datos de los participantes en múltiples condiciones o grupos experimentales y analizar los datos, sin modelar estas condiciones y grupos, un problema conocido como "pseudoreplicación"⁴⁴. Ejemplos de tales transgresiones incluyen la agrupación de datos para 10 participantes en tres condiciones y luego calcular los coeficientes de correlación entre las variables medidas como $n = 30$ en lugar de $n = 10$. Este enfoque rompe el supuesto de independencia de los casos, aumenta los grados de libertad y puede ser engañoso: de hecho, las correlaciones entre sujetos entre dos variables medidas a lo largo del tiempo pueden ser diferentes de las correlaciones entre sujetos entre dos variables medidas en el mismo punto en el tiempo.

15 Resultados

15a Los datos deben informarse de acuerdo con el Sistema Internacional de Unidades (es decir, unidades SI) o derivados aceptados. La excepción, como se aplica comúnmente en la ciencia de la nutrición, son las unidades para medir la ingesta o el gasto de energía, donde se puede usar kcal o kJ. Las medidas de tendencia central y variabilidad, utilizadas para describir cada conjunto de datos, deben ser claras, indicándose correctamente los efectos y la precisión estimada (por ejemplo, intervalo de confianza al 95%). Si se deben informar los valores p , se debe indicar que aquellos valores cercanos pero superiores a 0.05, no indican una "tendencia" o una aproximación a la significación estadística, mientras que un valor igual a cero no es posible⁴¹. Las unidades de medida no deben usarse en lugar de variables al describir los datos: por ejemplo, los datos deben denominarse ingesta de energía en lugar de ingesta de kJ (o calorías).

15b Siempre que sea razonable, debe hacerse un esfuerzo para presentar el rango completo de datos observados (es decir, mediciones o respuestas individuales) para ilustrar la consistencia de los efectos en lugar de solo estadísticas de resumen grupal. Cualquier normalización de las mediciones de acuerdo con otra variable o punto temporal (p. ej., % De cambio o g/kg) requiere una cuidadosa consideración y justificación, idealmente complementada por algún informe de los datos originales para que el lector pueda interpretar los hallazgos dentro del contexto de valores no ajustados/absolutos. Sin embargo, los participantes que producen un resultado que difiere en la dirección o la magnitud de los resultados medios del grupo no deben etiquetarse como "no respondedores" al tratamiento, basándose únicamente en esta observación aislada. Cuando se anticipa la probabilidad de variabilidad individual en respuesta a un tratamiento, el diseño del estudio podría incluir controles tales como pruebas repetidas del mismo tratamiento o la recopilación mecánica de datos que puedan corroborar diferencias reales en las respuestas fisiológicas/psicológicas (es decir, muestras biológicas que pueden verificar o explicar la eficacia del efecto). Se debe considerar la posibilidad de que la variabilidad diaria, en los resultados de la variable principal, explique resultados divergentes.

15c En aras de la divulgación de informes equilibrados, cualquier consecuencia negativa imprevista, daño o consecuencia no intencional, debe divulgarse completamente junto con los resultados principales (en lugar de incluirse en la Discusión). Por ejemplo, los problemas gastrointestinales inesperados asociados con un suplemento nutricional deben informarse en su totalidad,

independientemente de si los resultados principales, como las respuestas metabólicas y ergogénicas, se afectaron negativamente.

Discusión

16a Interpretación: la Discusión debe mantener el foco en los nuevos datos generados por la investigación, en lugar de revisar ampliamente la bibliografía general, aunque debe quedar claro cómo, los nuevos hallazgos proporcionados por la investigación, complementan y mejoran la evidencia existente. Las sugerencias más especulativas, que van más allá de los datos, pueden ser aceptables en esta sección, pero deben identificarse claramente como tales. Según el Título y la Introducción, nuevamente se debe tener especial cuidado en esta sección para garantizar que la elección del lenguaje evite la hipérbole o la inferencia incorrecta de la causalidad.

16b Generalización: se debe realizar consideraciones de cómo se pueden generalizar adecuadamente los resultados del proyecto de investigación, más allá de la población y del contexto en el que se midieron. El hecho de que la investigación involucre una muestra relativamente homogénea o heterogénea, no es una fortaleza ni una limitación en sí misma, pero esto tiene implicaciones sobre cómo los hallazgos, de manera amplia o específica, pueden trasladarse a otros.

16c Fortalezas y limitaciones: un resumen sucinto y veraz de las fortalezas del diseño, los procedimientos y los resultados del estudio, debe integrarse a lo largo de la discusión. Esto no solo brinda al lector (y revisor) la oportunidad de reflexionar sobre la confianza que depositan en los hallazgos del estudio, sino que, al llamar la atención sobre las cuidadosas características metodológicas y de comunicación de resultados, se alentará a otros investigadores a adoptar tales medidas en sus futuros trabajos. Las posibles fuentes de error o las variables confundentes también se deben indicar, cuando sea relevante. a lo largo de la Discusión, permitiendo que las conclusiones se atemperen en consecuencia. Por lo general, es preferible una referencia más integrada a las limitaciones que un párrafo independiente hacia el final de la Discusión, que enumere una selección de debilidades, que no se reflejan en la interpretación general. En cualquier caso, las conclusiones finales deben reflejar un equilibrio de todas las características señaladas en esta sección.

Otros

17 Conflicto de Intereses

Los autores tienen la responsabilidad de indicar aquellas relaciones que crean que otros podrían considerar que suponen conflictos de intereses o que deberían de divulgarse por cualquier otro motivo. Esto podría incluir, entre otros: relaciones con la industria, dependencia laboral/posesión de acciones, suministro de suplementos nutricionales o consumibles, venta de libros o cualquier otro sesgo por interés personal.

18 Protocolo

Se debe hacer referencia a cualquier protocolo publicado o registrado públicamente, y los cambios de cualquier registro anterior se deben enumerar y explicar.

Bibliografía

1. [Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *Trials*. 2010;11:32.](#)

2. [Dwan K, Li T, Altman DG, Elbourne D. CONSORT 2010 statement: extension to randomised crossover trials. *BMJ*. 2019;366,14378.](#)
3. [Byrnes WC, Clarkson PM, White JS, Hsieh SS, Frykman PN, Maughan RJ. Delayed onset muscle soreness following repeated bouts of downhill running. *J Appl Physiol*. 1985;59\(3\):710-5.](#)
4. [McHugh MP, Connolly DAJ, Eston RG, Gleim GW. Exercise-induced muscle damage and potential mechanisms for the repeated bout effect. *Sports Med*. 1999;27\(3\):157-70.](#)
5. [Maughan RJ. Returning to the writing of abstracts. *J Sports Sci*. 2004;22\(7\):603.](#)
6. [Brown AW, Bohan Brown MM, Allison DB. Belief beyond the evidence: using the proposed effect of breakfast on obesity to show 2 practices that distort scientific evidence. *Am J Clin Nutr*. 2013;98\(5\):1298-308.](#)
7. [Harriss DJ, Macsween A, Atkinson G. Standards for Ethics in Sport and Exercise Science Research: 2018 Update. *Int J Sports Med*. 2017;38\(14\):1126-31.](#)
8. [Winter EM, Maughan RJ. Requirements for ethics approvals. *J Sports Sci*. 2009;27\(10\):985.](#)
9. [Kordi R, Mansournia MA, Rostami M, Maffulli N. Troublesome decimals; a hidden problem in the sports medicine literature. *Scand J Med Sci Sports*. 2011;21\(3\):335-6.](#)
10. [Batterham AM, Atkinson G. How big does my sample size need to be? A primer on the murky world of sample size estimation. *Phys The Sport*. 2005;6\(3\):153-63.](#)
11. [Assel M, Sjoberg D, Elders A, Wang X, Huo D, Botchway A, et al. Guidelines for Reporting of Statistics for Clinical Research in Urology. *J Urol*. 2019;201\(3\):595-604.](#)
12. [Greenland S. Valid p-values behave exactly as they should:some misleading criticisms of p-values and their resolution with s-values. *Am Statistician*. 2019;73\(1\):106-14.](#)
13. [Bailey DM, Williams C, Betts JA, Thompson D, Hurst TL. Oxidative stress, inflammation and recovery of muscle function after damaging exercise: effect of 6-week mixed antioxidant supplementation. *Eur J Appl Physiol Physiol*. 2011;111\(6\):925-36.](#)
14. [Senn S, Julious S. Measurement in clinical trials: a neglected issue for statisticians? *Stat. Med.*, 2009;28\(26\):3189-209.](#)
15. [Currell K, Jeukendrup AE. Validity, reliability and sensitivity of measures of sporting performance. *Sports Med*. 2008;38\(4\):297-316.](#)
16. [Atkinson G, Nevill AM. Statistical methods for assessing measurement error \(reliability\) in variables relevant to sports medicine. *Sports Med*. 1998;26\(4\):217-38.](#)
17. [de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health Qual Life Outc*. 2006;4:54.](#)
18. [Cook JA, Hislop J, Adewuyi TE, Harrild K, Altman DG, Ramsay CR, et al. Assessing methods to specify the target difference for a randomised controlled trial: DELTA \(Difference ELicitation in TriAls\) review. *Health Technol Assess*. 2014;18\(28\):v-vi,1-175.](#)
19. [Cook JA, Julious SA, Sones W, Hampson LV, Hewitt C, Berlin JA, et al. DELTA\(2\) guidance on choosing the target difference and undertaking and reporting the sample size calculation for a randomised controlled trial. *BMJ*. 2018;363:k3750.](#)
20. [Schulz KF, Grimes DA. \(a\). Allocation concealment in randomised trials: defending against deciphering. *Lancet*. 2002;359\(9306\):614-8.](#)
21. [Schulz KF, Grimes DA. \(c\). Generation of allocation sequences in randomised trials: chance, not choice. *Lancet*. 2002;359\(9305\):515-9.](#)
22. [Schulz KF, Grimes DA. \(e\). Unequal group sizes in randomised trials: guarding against guessing. *Lancet*. 2002;359\(9310\):966-70.](#)
23. [Bernard C. An Introduction to the Study of Experimental Medicine. London: Courier Corporation. 1865](#)
24. [Schulz KF, Grimes DA. \(b\). Blinding in randomised trials: hiding who got what. *Lancet*. 2002;359\(9307\):696-700.](#)
25. [Funnell MP, Mears SA, Bergin-Taylor K, James LJ. Blinded and unblinded hypohydration similarly impair cycling time trial performance in the heat in trained cyclists. *J Appl Physiol* \(1985\). 2019;126\(4\):870-9.](#)
26. [Jeacocke NA, Burke LM. Methods to standardize dietary intake before performance testing. *Int J Sport Nutr Exerc Metab*. 2010;20\(2\):87-103.](#)
27. [Burke LM, Peeling P. Methodologies for Investigating Performance Changes With Supplement Use. *Int J Sport Nutr Exerc Metab*. 2018;28\(2\):159-69.](#)
28. [Wellek S, Blettner M. On the proper use of the crossover design in clinical trials: part 18 of a series on evaluation of scientific publications. *Dtsch Arztebl Int*. 2012;109, 276-281.](#)
29. [Bland JM, Altman DG. Best \(but oft forgotten\) practices: testing for treatment effects in randomized trials by separate analyses of changes from baseline in each group is a misleading approach. *Am J Clin Nutr*. 2015;102\(5\):991-4.](#)
30. [Thomas DM, Clark N, Turner D, Siu C, Halliday TM, Hannon BA, et al. Best \(but oft-forgotten\) practices: identifying and accounting for regression to the mean in nutrition and obesity research. *Am J Clin Nutr*. 2019;111\(2\):256-65.](#)
31. [Atkinson G, Batterham A. M. True and false interindividual differences in the physiological response to an intervention. *Exp Physiol*. 2015;100\(6\):577-88.](#)
32. [Bland JM, Altman DG. Transforming data. *BMJ*. 1996;312\(7033\):770.](#)
33. [Vickers AJ. Analysis of variance is easily misapplied in the analysis of randomized trials: a critique and discussion of alternative statistical approaches. *Psychosom Med*. 2005;67\(4\):652-5.](#)
34. [Vickers AJ. The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: a simulation study. *BMC Med Res Methodol*. 2001;1:6.](#)
35. [Bland JM, Altman DG. \(a\) Calculating correlation coefficients with repeated observations: Part 1--Correlation within subjects. *BMJ*. 1995;310\(6977\):446.](#)
36. [Bland JM, Altman DG. \(b\). Calculating correlation coefficients with repeated observations: Part 2--Correlation between subjects. *BMJ*. 1995;310\(6980\):633.](#)
37. [Atkinson G, Williamson P, Batterham AM. Issues in the determination of 'responders' and 'non-responders' in physiological research. *Exp Physiol*. 2019;104\(8\):1215-25.](#)
38. [Hecksteden A, Kraushaar J, Scharhag-Rosenberger F, Theisen D, Senn S, Meyer T. Individual response to exercise training - a statistical perspective. *J Appl Physiol* \(1985\). 2015;118\(12\):1450-9.](#)
39. [Lonergan M, Senn SJ, McNamee C, Daly AK, Sutton R, Hattersley A, et al. Defining drug response for stratified medicine. *Drug Discov Today*. 2017;22\(1\):173-9.](#)
40. [Bland JM, Altman DG. Some examples of regression towards the mean. *BMJ*. 1994;309\(6957\):780.](#)
41. [Tu YK, Gilthorpe MS. Revisiting the relation between change and initial value: a review and evaluation. *Stat Med*. 2007;26\(2\):443-57.](#)
42. [Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol*. 2016;31\(4\):337-50.](#)
43. [Schulz KF, Grimes DA. \(d\). Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet*. 2002;359\(9308\):781-5.](#)
44. [Lazic SE. The problem of pseudoreplication in neuroscientific studies: is it affecting your analysis? *BMC Neurosci*. 2010;11:5.](#)

ANEXO

PRESENT 2020 (INFORME ADECUADO DE LA EVIDENCIA EN ENSAYOS CLÍNICOS DE NUTRICIÓN DEL DEPORTE Y EL EJERCICIO)

LISTA DE VERIFICACIÓN DE LA INFORMACIÓN A INCLUIR AL INFORMAR SOBRE INVESTIGACIÓN EN

NUTRICIÓN DEPORTIVA Y METABOLISMO DEL EJERCICIO*

© 2020 Human Kinetics, Inc.

| Sección | Ítem | Lista de Verificación | Página/Número de línea (Si es aplicable) |
|-----------------------------------|------|---|---|
| Título | | | |
| | 1a | Indica las variables independientes (grupos / condiciones) y dependientes (resultado) | |
| | 1b | Identifica la población de estudio o el caso. | |
| Resumen | | | |
| | 2a | Especifica el diseño de investigación, los métodos y las características de la población de estudio. | |
| | 2b | Informa de manera equilibrada los resultados y cita los datos actuales | |
| | 2c | Las conclusiones se ciñen a las variables medidas, sin especulaciones ni recomendaciones no compatibles. | |
| Introducción | | | |
| | 3a | Presenta una justificación científica basada en una revisión objetiva de la evidencia disponible. | |
| | 3b | Indica los propósitos, objetivos, preguntas de investigación y/o hipótesis | |
| Métodos | | | |
| <i>Ética</i> | 4 | Proporciona detalles de aprobación ética (citando conducta de investigación en humanos de acuerdo con la Declaración de Helsinki) | |
| <i>Diseño</i> | 5 | Resume el diseño de la investigación (por ejemplo, ensayo paralelo/cruzado, aleatorizado, contrabalanceado, cegamiento, observacional) | |
| <i>Muestreo</i> | 6a | Enumera los criterios de elegibilidad (inclusión/exclusión) y el método de muestreo | |
| | 6b | Caracteriza la muestra del estudio (por ejemplo, demografía, antropometría, estilo de vida) | |
| | 6c | Informa del entorno/ubicación y sobre los períodos de reclutamiento y recogida de datos. | |
| | 6d | Justifica el tamaño de la muestra (presentando el tamaño del efecto objetivo seleccionado y las variaciones de error para replicar las estimaciones del tamaño de la muestra) | |
| <i>Intervenciones[‡]</i> | 7 | Detalla todos los aspectos de los grupos/condiciones (considerando la necesidad de verificar la composición de las sustancias ingeridas) | |
| <i>Mediciones</i> | 8a | Define las prestablecidas variables de resultado principales, secundarias y/o mecánicas | |
| | 8b | Racionaliza la selección de protocolos de prueba, considerando la validez y la confiabilidad (por ejemplo, coeficiente de variación, familiarización) | |
| | 8c | Justifica el mínimo efecto relevante o la mínima diferencia clínicamente relevante | |
| <i>Aleatorización</i> | 9 | Detalla los mecanismos concretos de generación y asignación y de cegamiento de la asignación aleatorizada | |
| <i>Cegamiento[‡]</i> | 10 | Documenta si los participantes y/o investigadores conocían la asignación (por ejemplo, cuestionario de salida) | |
| <i>Estandarización</i> | 11 | Describe los controles dentro y entre los participantes (por ejemplo, replicación/informe de dieta, actividad física, sueño, ciclo menstrual) | |
| <i>Efectos de orden</i> | 12 | Controla en detalle las influencias sistemáticas de las mediciones en serie (por ejemplo, efecto de secuencia en el modelo de análisis, intervalo de lavado) | |

(continúa)

Lista de verificación PRESENT 2020 (continúa)

© 2020 Human Kinetics, Inc.

| Sección | Ítem | Lista de Verificación | Página/Número de línea (Si es aplicable) |
|-------------------------------|------|--|---|
| <i>Estadísticas</i> | 13a | Especifica el contraste para las inferencias primarias (es decir, en relación con el control apropiado, no los cambios desde el inicio en cada grupo/condición) | |
| | 13b | Distingue claramente y justifica por completo cualquier análisis de subgrupos no planificado, provisional o exploratorio | |
| | 13c | Describe cualquier ajuste por supuestos estadísticos inadecuados y por covariables relevantes (por ejemplo, medidas de referencia) | |
| Resultados | | | |
| <i>Flujo de Participantes</i> | 14a | Informa del tamaño de la muestra en cada fase desde el reclutamiento hasta el análisis (con motivos de pérdidas y exclusiones) | |
| | 14b | Asegura que el análisis de datos coincida con el diseño de la investigación, evitando la agrupación de datos entre grupos/condiciones (es decir, pseudoreplicación) | |
| <i>Resultados</i> | 15a | Informa utilizando las unidades del Sistema Internacional e informa sobre las medidas de tendencia central, variabilidad y tamaño/precisión del efecto (intervalos de confianza) | |
| | 15b | Aporta datos/respuestas individuales (por ejemplo, dibuja figuras que muestran los datos en bruto en cada grupo/condición) | |
| | 15c | Documenta todos los daños relevantes y consecuencias no deseadas observadas | |
| Discusión | | | |
| | 16a | Presenta una interpretación objetiva y equilibrada de los datos observados dentro del contexto de la evidencia existente. | |
| | 16b | Considera la aplicabilidad y/o relevancia práctica de los resultados de la investigación (por ejemplo, validez externa) | |
| | 16c | Reconoce las fortalezas y limitaciones de la investigación relevantes para la interpretación adecuada (por ejemplo, validez interna) | |
| Otros | | | |
| <i>Conflicto Intereses</i> | 17 | Indica cualquier relación relevante (por ejemplo, soporte financiero, técnico, material) | |
| <i>Protocolo</i> | 18 | Identifica cualquier protocolo registrado público o publicado (explicando cualquier desviación) | |

*Adaptado de la lista de verificación CONSORT 2010 para comunicar resultados de ensayos controlados aleatorios y puede usarse junto con el documento asociado que se expande en cada ítem.

‡Los ítems 7 (Intervenciones) y 10 (Cegamiento) son relevantes para la investigación experimental, incluidos los contrastes simples/dobles ciegos de suplementos nutricionales.