

Número necesario a tratar y número necesario para producir daño: su importancia

Number Needed to Treat and Number Needed to Harm: Its Importance

Recibido: 11 Noviembre 2020 | Aceptado: 02 Febrero 2021

MARÍA PAULA JASSIR ACOSTA^a

Médica, Departamento de Epidemiología Clínica y Bioestadística, Pontificia Universidad Javeriana, Bogotá, Colombia

ORCID: <https://orcid.org/0000-0002-9682-2759>

CARLOS GÓMEZ-RESTREPO

Médico, MSc, PhD. Decano de la Facultad de Medicina, Pontificia Universidad Javeriana, Bogotá, Colombia

ORCID: <https://orcid.org/0000-0002-9013-5384>

RESUMEN

Para lograr una adecuada comprensión de los resultados de un estudio de investigación en salud es importante tener claros ciertos conceptos epidemiológicos y estadísticos. El objetivo de este artículo es abordar, de manera sencilla y clara, los conceptos de número necesario a tratar y número necesario para producir daño, explicar cómo se calculan, aclarar la importancia de su interpretación y aplicación en contextos clínicos a partir de ejemplos, al igual que algunas de las limitaciones de su uso que han sido frecuentemente discutidas en la literatura.

Palabras clave

número necesario a tratar; epidemiología; bioestadística.

ABSTRACT

In order to comprehend the results of research studies in health it is important to have clarity about certain epidemiologic and statistical concepts. The objective of this article is to explain, in a clear and simple manner, the concepts of the number needed to treat and number needed to harm, how to calculate these measures, the importance of the interpretations of their values and its application in clinical contexts with examples. We will also talk about the limitations of their use that have been frequently discussed in the literature.

Keywords

number needed to treat; epidemiology; biostatistics.

^a Autora de correspondencia: m.jassir@javeriana.edu.co

Cómo citar: Jassir Acosta MP, Gómez-Restrepo C. Número necesario a tratar y número necesario para producir daño: su importancia. Univ. Med. 2021;62(2). <https://doi.org/10.11144/Javeriana.umed62-2.nnnt>

Introducción

Existen muchas mediciones en el área de la salud que se derivan de los diferentes reportes obtenidos a partir de la aplicación de diversos diseños de investigación. Entre estos tenemos: los riesgos relativos, la reducción absoluta de riesgo (RAR), los riesgos relativos indirectos, las razones de momios o, como comúnmente se les conoce, *odds ratio* (OR), entre

otras (1). Estas mediciones aportan a las decisiones en salud y en salud pública para los tomadores de decisiones, aparte de que proveen herramientas que pueden servir en la toma de decisiones (2). Unas medidas más recientes son el número necesario a tratar (NNT) y el número necesario para producir daño (NNH), que son la base de este trabajo.

El concepto de NNT lo introdujeron en la literatura médica por primera vez, en 1998, Laupacis et al. (1), e inicialmente fue considerado una medida útil para la interpretación de ensayos clínicos aleatorizados, utilizados habitualmente para evaluar la seguridad, la eficacia y los posibles efectos de tratamientos o intervenciones en salud (3). Sin embargo, a través de los años, el NNT también se ha utilizado para expresar las diferencias entre tratamientos como resultado de otros tipos de estudios comparativos, como las revisiones sistemáticas, los metanálisis y algunos estudios analíticos observacionales (casos y controles, cohortes) (2). Desde su introducción, el NNT ha sido muy útil para los médicos en la toma de decisiones en salud, en aras de asegurarles a sus pacientes un tratamiento adecuado y basado en la evidencia, igual que para los tomadores de decisiones en relación con la regulación de medicamentos (2).

No obstante, para interpretar de modo acertado los resultados de estudios epidemiológicos es necesario tener claras las distintas medidas con las cuales se expresan sus resultados. Aun cuando existen varias medidas de impacto implementadas en los distintos tipos de estudios epidemiológicos, según su relevancia y uso en la práctica médica, en este escrito abordamos específicamente el NNT y el NNH.

Número necesario a tratar

El NNT se define como el número de pacientes que es necesario tratar para evitar un suceso indeseable o tener un desenlace favorable determinado (2, 4). El NNT es una estadística derivada, calculada a partir de la incidencia acumulada de un evento observado en sujetos a través del tiempo, posterior a una intervención

o tratamiento (5). La incidencia acumulada se entiende como la probabilidad o el riesgo medio de los miembros de una población de desarrollar un evento en un tiempo determinado y se calcula mediante la división del número de casos nuevos de un evento en un periodo sobre el total de la población al inicio de la evaluación (5). Este cálculo se evidencia a continuación:

$$\text{Incidencia Acumulada (IA)} = \frac{\text{Número de casos nuevos de un evento}}{\text{Número total de la población al inicio}}$$

El NNT se calcula y se entiende como el inverso de la RAR, otra medida de impacto que estima la disminución real de desenlaces desfavorables en pacientes que reciben un tratamiento en particular (6). La RAR se calcula a partir de la diferencia entre la incidencia acumulada del grupo tratado con la intervención y la incidencia acumulada del grupo placebo, y expresa las consecuencias de dar un tratamiento específico (6). En la tabla 1, presentamos las fórmulas para calcular el NNT y la RAR.

Tabla 1
Fórmulas para calcular NNT y RAR

Medida	Fórmula
NNT	1/RAR
RAR	IAc - IAi

IAc: incidencia acumulada en el grupo control; IAi: incidencia acumulada en grupo de la intervención.

Número necesario para producir daño

El NNH, al igual que el NNT, es una estadística derivada, calculada a partir de la incidencia acumulada de efectos adversos o de eventos perjudiciales (7). Esta medida expresa cuántos pacientes se necesitan tratar con un tratamiento o intervención para que un paciente adicional presente un efecto adverso específico (7). El NNH, al igual que el NNT, se calcula a partir de un inverso (tabla 2); pero en este caso se utiliza el aumento absoluto del riesgo (AAR), que corresponde a la diferencia entre la incidencia acumulada o el riesgo de la población tratada con

una intervención y el de la población no tratada (6).

Tabla 2
Fórmulas para calcular NNH y AAR

Medida	Fórmula
NNH	1/AAR
AAR	IAi - IAc

IAC: incidencia acumulada en grupo control; IAi: incidencia acumulada en grupo de la intervención

Intervalos de confianza

La RAR, al igual que la AAR, se considera un punto estimado y su valor se encuentra acompañado de un rango de valores que calculan un parámetro de certeza alrededor de un punto, es decir, dentro de un intervalo de confianza (8, 9). El intervalo de confianza presenta un rango de valores que son mayores y menores al punto estimado en cuestión (10). Un intervalo de confianza del 95 % indica que dentro del rango de valores se encuentra el valor real del parámetro con un 95 % de certeza (9).

El cálculo del intervalo de confianza se basa en una fórmula que tiene en cuenta el error estándar, un indicador que mide el error aleatorio en un dato estadístico informado (11). Aun cuando el cálculo del error estándar se basa en una fórmula, también se puede generar de manera automática mediante la utilización de herramientas habilitadas en internet, como calculadoras específicas para los intervalos de confianza (12). Dada su complejidad matemática, en este artículo no se ahondará en la fórmula específica del error estándar, pero se menciona, ya que este va incluido en la fórmula simple para el cálculo del intervalo de confianza de la RAR (13), que se encuentra expresada a continuación:

IC 95 % RAR = RAR ± (1,96 × error estándar de RAR)

Aun cuando existen varios métodos matemáticos implementados para abordar el cálculo del intervalo de confianza, todavía hay ciertas limitaciones en su uso. Este tema se aborda más adelante.

Teniendo en cuenta lo anterior, al expresar la RAR es común que el NNT venga acompañado de su propio intervalo de confianza, ya que ambos son parámetros estadísticos sujetos a incertidumbre(13). En este caso, el intervalo de confianza del NNT corresponde asimismo a los inversos del intervalo de confianza de la RAR (8). Se debe tener en cuenta que los IC95 % dependen del tamaño de la muestra, es decir, que entre menos pacientes tenga la muestra de un estudio, el IC95 % será más amplio y por esto, los valores límite del IC95 % también se deben tener presente al momento de interpretar los resultados (14).

Interpretación y Cálculo del NNT y NNH

Comúnmente, el NNT se interpreta como la cantidad de pacientes que se necesitan tratar para evitar un suceso desfavorable o para que un paciente adicional obtenga un resultado favorable, como respuesta a dicha intervención o tratamiento en comparación con el grupo control u otro tratamiento (2, 7). Es importante tener en cuenta que para una interpretación correcta del NNT, o el NNH, siempre se debe mencionar el comparador, ya sea otro tratamiento o placebo, con el que se está realizando el análisis (6).

Abordaremos la interpretación del NNT con un ejemplo hipotético de un ensayo clínico aleatorizado en el cual se investigó la eficacia de un nuevo medicamento X para la prevención de infartos agudos de miocardio (IAM) en mujeres posmenopáusicas (4). Del total de 800 mujeres en el estudio, 400 fueron aleatorizadas y asignadas al grupo que recibió placebo, y 400, al grupo que recibió el medicamento X. De las 400 mujeres del grupo placebo, 30 presentaron IAM y de las 400 mujeres del grupo experimental, aquellas que recibieron el medicamento X, 10 presentaron IAM. Presentamos los resultados en la tabla 3.

Tabla 3
Resultados de ensayo clínico aleatorizado hipotético

	IAM	Sin IAM	Total
Medicamento X	10	390	400
Placebo	30	370	400

IAM: infarto agudo de miocardio.

Partiendo del caso hipotético y de los conceptos revisados, podemos calcular el NNT para la prevención de IAM del medicamento X comparado con el placebo:

$$IA_i = 10/400 = 0,025$$

$$IA_c = 30/400 = 0,075$$

$RAR = 0,075 - 0,025 = 0,05$ **Error estándar de la RAR = 0,015**

$IC\ 95\ \% (RAR) = RAR \pm (1,96 \times \text{error estándar de RAR})$

$$IC\ 95\ \% (RAR) = 0,05 \pm (0,0294)$$

$$IC95\ \% (RAR) = (0,02 - 0,079)$$

$$NNT = 1/0,05 = 20$$

$$IC95\ \% NNT = (13 \text{ a } 50)$$

En el ejemplo anterior, vemos que para evitar un IAM en mujeres posmenopáusicas sería necesario tratar a 20 mujeres con el medicamento X con un intervalo de confianza de 13 a 50. Esto significa que la variación debida al azar sugiere que el NNT podría ser tan bueno como 13 o como 50 pacientes necesarios a tratar (4). Aclaremos que usualmente el NNT, y asimismo el NNH, se redondea al número entero más cercano, pues este expresa un número de pacientes y no sería posible una adecuada interpretación con un número fraccionario (14).

No obstante, es importante entender la naturaleza del desenlace o evento que se está estudiando, es decir, si el evento es un desenlace favorable o desfavorable, al interpretar los resultados de un estudio (15). Lo anterior es relevante cuando el NNT nos da un valor negativo. Cuando el NNT da un valor negativo, se podría hablar de dos escenarios: la intervención hace daño o es una intervención con desenlaces positivos. Cuando se está hablando de evitar un evento desfavorable en torno a un tratamiento o intervención, como en el ejemplo anterior, un NNT negativo significaría que dicho tratamiento o intervención es perjudicial (6). Contrariamente, si se está estudiando un evento favorable, un

NNT negativo indicaría que el tratamiento o intervención es beneficiosa. Conceptualmente, se podría decir que el NNT negativo es el equivalente al NNH (6); sin embargo, cuando se habla de desenlaces positivos en términos de un NNH, puede haber confusión.

Para explicar lo anterior, tomaremos como ejemplo la interpretación de los resultados de un ensayo clínico aleatorizado, doble ciego, que estudió la efectividad del aripiprazol lauroxil versus placebo para la mejoría de los síntomas positivos y negativos de la esquizofrenia, definido en el estudio como una mejoría igual o mayor al 30 % de la puntuación basal de la Escala de los Síndromes Positivo y Negativo, en pacientes que cursan con una exacerbación aguda de esquizofrenia (16). Los autores encontraron que la incidencia acumulada de la mejoría en los pacientes tratados con aripiprazol lauroxil y placebo fue del 35,3 % y del 18,4 %, respectivamente. En este caso, el cálculo de la RAR sería el siguiente:

$$RAR = 0,184 - 0,353 = -0,169$$

$$NNT = -6$$

Sin embargo, los autores han realizado la resta inversamente, es decir, calculando el AAR:

$$AAR = 0,353 - 0,184 = 0,169$$

$$NNH = 6$$

Al convertir la RAR o la AAR a sus respectivos inversos, los resultados son -6 y 6 , respectivamente. Sin embargo, lo importante en este caso es la manera como se aborda la interpretación del resultado. En este caso, se está haciendo referencia a cuántos pacientes hay que tratar para que haya una mejoría, es decir, para que suceda un desenlace beneficioso. Por lo tanto, se toma el AAR como el aumento absoluto del riesgo de que haya un desenlace beneficioso. Teniendo en cuenta que el NNH hace referencia a desenlaces desfavorables y factores de riesgo; por lo tanto, es confuso hablar de NNH en términos de un desenlaces favorables o beneficiosos. Los autores en el artículo lo abordan desde el NNT y afirman que para la presentación de mejorías en la exacerbaciones de la esquizofrenia con aripiprazol lauroxil en comparación con el placebo el NNT de 6 (16). No obstante, lo que cabe resaltar del ejemplo es

que se debe partir del desenlace buscado para poder interpretar las medidas de modo adecuado.

Ejemplos de la interpretación del NNT en la práctica clínica

En comparación con placebo

Como otro ejemplo, tomaremos el cálculo del NNT en un estudio que compara la efectividad de la oxcarbacepina versus placebo para reducir el dolor neuropático del nervio periférico (17). De los 83 pacientes incluidos en el estudio, 16 tuvieron una disminución significativa del dolor, definido este el estudio como una reducción igual o mayor del 50 % (17), con la administración de oxcarbacepina y 4 pacientes tuvieron la misma respuesta con el placebo. En este caso, las incidencias acumuladas de la oxcarbacepina versus el placebo serían del 19,27 % y del 4,81 %, respectivamente. Esto significa que la oxcarbacepina, comparada con el placebo, causaría una respuesta hacia la disminución significativa del dolor en el 14,46 % de los pacientes, es decir, que la RAR en este caso sería de un -14,46 %. Similarmente, vemos que se encuentra un NNT negativo, ya que nos encontramos en el contexto del riesgo de presentar un desenlace positivo, en este caso, disminución significativa del dolor periférico. En este caso, las incidencias acumuladas serían:

$$IA_i = 16/83 = 0,1927$$

$$IA_c = 4/83 = 0,0481$$

Entonces, si $RAR = IA_c - IA_i$:

$$RAR = 0,0481 - 0,1927 = -0,1446$$

$$NNT = -6,9 = -7$$

Ahora bien, ¿cómo interpretamos este número? En este caso, un NNT de -7 para un desenlace positivo significaría que debo tratar a 7 pacientes para que uno presente reducción del dolor. En el artículo, los autores prescinden del signo negativo, posiblemente por la posibilidad de confusión; sin embargo, al hacer el ejercicio, se evidencia que observando el valor, esto es, las incidencias acumuladas para un desenlace positivo, tiene sentido que el resultado a favor de la intervención en términos de desenlaces

positivos sea un número negativo, ya que se esperaría que la incidencia acumulada del grupo de intervención fuese mayor a la del grupo control.

En comparación con otras líneas de tratamiento

Anteriormente habíamos tomado ejemplos de comparación entre NNT o NNH y un placebo. Sin embargo, es importante tener en cuenta que, asimismo, se podrían realizar interpretaciones de estas medidas entre distintos tratamientos de elección (18). Esto es particularmente importante en decisiones de la práctica clínica.

Para esto, tomaremos como ejemplo los resultados del uso de la cariprazina versus placebo para el tratamiento de la esquizofrenia en una revisión de ensayos clínicos aleatorizados (18). Uno de los efectos adversos más comunes ante la administración de cariprazina es la acatisia con un NNH de 12 a dosis altas (4,5-6 mg/dl). Esto significa que al administrarle cariprazina a 12 pacientes, uno de ellos presentaría acatisia como un efecto adverso. Los autores mencionan que al comparar el NNH de la cariprazina con el de otros antipsicóticos, puntualmente el aripiprazol y la lurasidona, para la presentación de acatisia como efecto adverso, encuentran que estos medicamentos tienen un NNH de 12, 25 y 10, respectivamente. Siendo así, entendemos que, en comparación con la cariprazina, en el caso de la administración de lurasidona, se necesitaría tratar a menos pacientes para producir acatisia y, de manera similar, con la administración de aripiprazol se necesitaría tratar a más pacientes para producir acatisia. Entender este tipo de ejemplos es fundamental a la hora discernir entre distintas líneas de tratamiento durante la práctica clínica, ya que resulta pertinente tener en cuenta las comparaciones de los NNH para efectos adversos entre las otras opciones de tratamiento (18, 19).

Limitaciones

Algunos autores consideran que la interpretación del NNT puede verse obstaculizada por las altas

incidencias acumuladas observadas al placebo en algunos ensayos clínicos aleatorizados (20). Sin embargo, ambas medidas resultan útiles al estratificar distintas intervenciones a partir de su eficacia y, asimismo, ayudaría a escoger entre diferentes líneas de tratamiento que en principio son similares, siempre y cuando se haga un análisis adecuado de contexto (20).

Por ejemplo, tomemos la situación hipotética de un estudio que compara la reducción del riesgo de presentar un evento adverso en específico entre un medicamento Y versus placebo. Las incidencias acumuladas de dicho evento adverso corresponden al 2 % y al 15 %, es decir, 0,02 y 0,15, respectivamente. Por lo tanto, tenemos que la RAR sería del 13 %, es decir, 0,13, y que el NNT correspondería a 8. Se podría concluir que se necesita tratar a 8 pacientes con el medicamento Y para evitar que se presente el evento adverso. Ahora bien, si las cifras de las incidencias acumuladas del mismo medicamento X versus el placebo fueran 80 % y 93 %, es decir, 0,8 y 0,93, tendríamos la misma RAR y el mismo NNT. Sin embargo, al hacer un análisis de ambas situaciones, encontramos que, en la primera, la incidencia acumulada al placebo es relativamente baja y probablemente sería beneficioso administrar el medicamento. En la segunda situación, las incidencias acumuladas, tanto del medicamento como del placebo, son muy altas y por esto podríamos considerar que si bien existe la misma diferencia en las incidencias acumuladas, es decir, la misma RAR y el mismo NNT, podríamos considerar innecesario administrar el medicamento, ya que el riesgo de presentar el evento adverso con el medicamento es tan alta como el 80 %.

De este último ejemplo se deriva otra problemática. Si se considera que desde su concepto inicial, y especialmente en ensayos clínicos aleatorizados, una de las ventajas de la utilización del NNT es que este expresa las consecuencias de prescribir un medicamento versus placebo, se está equiparando el placebo a la no administración de tratamiento (21). Sin embargo, cierta evidencia muestra que el efecto placebo podría inducir, en algunos casos, efectos terapéuticos deseados; por lo tanto, sería

incorrecto afirmar que cuando se trata de hacer un análisis comparativo entre un tratamiento específico y el placebo, administrar el placebo sería equivalente a no dar tratamiento (21). Consecuentemente, se debe entender que en el caso del NNT, la cifra por sí sola no proporciona el número de pacientes que tendrían una respuesta favorable por medio de los mecanismos placebo o aquellos que no mostrarían respuesta en lo absoluto (7). Para tener en cuenta esas cifras, se debe volver a los datos a partir de los cuales se realizó el cálculo del NNT.

Otro tema para tener en cuenta es la interpretación y metodología para calcular los IC95 % del NNT y NNH. De los métodos matemáticos que existen, el método Wald simple ha sido el más utilizado (22). Sin embargo, en ciertas situaciones este método podría generar resultados poco confiables o hasta teóricamente imposibles (22). Lo anterior sucede cuando, por ejemplo, el IC95 % del RAR abarca el 0, dado que el IC95 % del NNT llegaría hasta el infinito y su interpretación sería compleja (4, 22). Adicionalmente, este método tiene otra limitante: los IC95 % que computa en casos de muestras pequeñas resultan ser muy estrechos, especialmente cuando la tasa de finalización es menor al 50 %, por lo que la cobertura del intervalo de confianza no sería 95 %, sino alrededor del 60 % en promedio (23).

Teniendo en cuenta las limitaciones de este método, se propuso la utilización de una versión ajustada, llamada *método Wald ajustado* y es actualmente el método preferido para calcular el IC95 % (23). Este computa el IC95 % mediante la suma de dos eventos exitosos y dos eventos fallidos a la incidencia acumulada observada, es decir, adicionar un 2 y un 4 al numerador y al denominador, respectivamente, y luego aplicar la fórmula Wald simple (23). Aun cuando existen otros métodos como alternativa para el cálculo de los IC95 %, algunos generan intervalos de confianza muy conservadores (22, 23) y todos tienen recomendaciones específicas para su uso, que usualmente dependen de variables como el tamaño de la muestra, si los límites del IC95 % se acercan a 1 o 0, entre otros (22, 24).

Por último, se ha planteado a favor del uso del NNT que esta es una medida que permite comparar 2 o más tratamientos valorados en distintos ensayos clínicos (4). Sin embargo, lo anterior ha sido debatido, dado que el valor del NNT varía según el tiempo de seguimiento de un estudio (4). Por lo tanto, sería erróneo comparar el NNT de dos ensayos clínicos cuyo tiempo de seguimiento es distinto (4). Otro factor para tener en cuenta son los riesgos basales de las poblaciones estudiadas, ya que podrían variar por distintos factores (genéticos, dietarios, de exposición, etc.) y, por lo tanto, no tendría validez contrastar entre poblaciones distintas (4). Esto último es pertinente cuando se pretende aplicar estas medidas en estudios como los metanálisis (14).

Conclusión

Desde su introducción a la literatura médica, el objetivo de las medidas de impacto discutidas ha sido permitir la comparación de riesgos y beneficios en aproximaciones preventivas, terapéuticas, diagnósticas y de rehabilitación en el ámbito de la salud (1). A través de los años, su uso ha servido de guía para la toma de decisiones en el ámbito clínico rutinario y de apoyo en evaluaciones riesgo-beneficio para la toma de decisiones frente a la regulación de medicamentos (2).

Actualmente, resulta de gran importancia tener claras las definiciones de NNT y NNH a la hora de interpretar los resultados del impacto de algún tipo de intervención o tratamiento en salud. Es importante también la manera cómo se comunican los resultados de un estudio, al igual que la interpretación que se le realiza, dado que influye decisivamente en la práctica clínica y decisiones en salud pública.

Si bien este artículo resalta la pertinencia del uso de estas medidas de impacto, se deben tener en cuenta las limitaciones de su cálculo, su interpretación y las discusiones acerca de la aplicabilidad de las medidas, teniendo en cuenta el contexto en cual que se esté aplicando. Los profesionales de la salud y los

tomadores de decisiones deben considerar todas las características y las limitaciones mencionadas, al igual que las características poblacionales e individuales de los pacientes, a la hora de tomar acciones tanto en salud pública como en el tratamiento de pacientes.

Conflictos de intereses

Los autores no tienen conflictos de intereses.

Referencias

1. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med.* 1998;318(26):1728-33. <https://doi.org/10.1056/NEJM198806303182605>
2. Mendes D, Alves C, Batel-Marques F. Number needed to treat (NNT) in clinical literature: an appraisal. *BMC Med.* 2017;15(1):112. <https://doi.org/10.1186/s12916-017-0875-8>
3. Lazcano-Ponce E, Salazar-Martínez E, Gutiérrez-Castrellón P, Ángeles-Llerenas A, Hernández-Garduño A, Viramontes JL. Ensayos clínicos aleatorizados: variantes, métodos de aleatorización, análisis, consideraciones éticas y regulación. *Salud Pública Mex.* 2004;46(6):559-84.
4. Gómez-Acebo I, Dierssen-Sotos T, Llorca J. Number needed to treat: interpretation and estimation in multivariable analyses and censored data. *Med Clin (Barc).* 2014;142(10):451-6. <https://doi.org/10.1016/j.medcli.2013.05.003>
5. Moreno-Altamirano A, López-Moreno S, Corcho-Berdugo A. Principales medidas en epidemiología. *Salud Pública Mex [internet].* 2015;42(4):338-48. Disponible en: <https://www.scielosp.org>

g/article/ssm/content/raw/?resource_ssm_path=/media/assets/spm/v42n4/2882.pdf

6. Otero W, Sierra F. Medidas de impacto o magnitud del beneficio de una intervención. *Rev Colomb Gastroenterol*. 2003;18(1):46-8.

7. Andrade C. The numbers needed to treat and harm (NNT, NNH) statistics: what they tell us and what they do not. *J Clin Psychiatry*. 2015;76(3):e330-3.

8. Osiri M, Suárez-Almazor ME, Wells GA, Robinson V, Tugwell P. Number needed to treat (NNT): implication in rheumatology clinical practice. *Ann Rheum Dis*. 2003;62(4):316-21. <https://doi.org/10.1136/ard.62.4.316>

9. Perry DC, Griffin XL, Dritsaki M, Costa ML, Parsons N. Becoming confident about confidence intervals. *Bone Joint J*. 2017;99(5):9-11. <https://doi.org/10.1302/0301-620X.99B5.BJJ-2017-0075>

10. Kamper SJ. Confidence intervals: linking evidence to practice. *J Orthop Sport Phys Ther*. 2019;49(10):763-4. <https://doi.org/10.2519/jospt.2019.0706>

11. Luko SN. El concepto del error estándar [internet]. *Standardization News*. 2012. p. 1. Disponible en: <https://sn.astm.org/?q=data-points/standard-error-concept-nd12.html>

12. Evidence based Medicine Toolbox. Randomized controlled trial calculator [internet]. [Cited 2021 ene 31]. Disponible en: <https://ebm-tools.knowledgetranslation.net/calculator/randomized/>

13. Molinero LM. Número de pacientes que será necesario tratar NNT. *Sociedad Española de Hipertensión*. 2004. p. 1-7.

14. DiCenso A. EBN notebook: clinically useful measures of the effects of treatment. *Evid Based Nurs*. 2001;4:36-9.

15. Baptista González H. El número necesario a tratar (NNT) y número necesario para hacer daño (NNH). *Rev Investig Médica Sur, México* [internet]. 2008;15(4):302-3. Disponible en: <http://www.medigraphic.com/pdfs/medsur/ms-2008/ms084h.pdf>

16. Citrome L, Du Y, Weiden PJ. Assessing effectiveness of aripiprazole lauroxil vs placebo for the treatment of schizophrenia using number needed to treat and number needed to harm. *Neuropsychiatr Dis Treat*. 2019;15:2639-46. <https://doi.org/10.2147/NDT.S207910>

17. Demant DT, Lund K, Vollert J, Maier C, Segerdahl M, Finnerup NB, et al. The effect of oxcarbazepine in peripheral neuropathic pain depends on pain phenotype: A randomised, double-blind, placebo-controlled phenotype-stratified study. *Pain*. 2014;155(11):2263-73. <http://dx.doi.org/10.1016/j.pain.2014.08.014>

18. Citrome L. Cariprazine for the treatment of schizophrenia: a review of this dopamine D3-Preferring D3/D2 receptor partial agonist. *Clin Schizophr Relat Psychoses*. 2016;10(2):109-19.

19. Barratt A, Wyer PC, Hatala R, McGinn T, Dans AL, Keitz S, et al. Tips for learners of evidence-based medicine: 1. Relative risk reduction, absolute risk reduction and number needed to treat. *Cmaj*. 2004;171(4):353-8.

20. Citrome L, Ketter TA. NNT and NNH remain helpful in evidence-based medicine. *Br J Psychiatry*. 2016;209(3):262-4.

21. Roose SP, Rutherford BR, Wall MM, Thase ME. Practising evidence-based medicine in an era of high placebo response: Number needed to treat reconsidered. *Br J Psychiatry*. 2016;208(5):416-20.

22. Bender R. Calculating confidence intervals for the number needed to treat. *Control Clin Trials*. 2001;22:102-10.
23. Sauro J, Lewis JR. Estimating completion rates from small samples using binomial confidence intervals: Comparisons and recommendations. *Proc Hum Factors Ergon Soc*. 2005;2100-4.
24. MeasuringU. Confidence interval calculator for a completion rate [internet]. [Citado 2020 oct 16]. Disponible en: <https://measuringu.com/wald/>