

## Número de pacientes que será Necesario Tratar NNT

Preparado por Luis M. Molinero (Alce Ingeniería)

CorreoE: [bioestadistica@alceingenieria.net](mailto:bioestadistica@alceingenieria.net)

[Artículo en formato PDF](#)

[www.seh-lelha.org/stat1.htm](http://www.seh-lelha.org/stat1.htm)

Marzo 2004

Existen diferentes parámetros para cuantificar el efecto de un tratamiento cuando el resultado se determina mediante una variable binaria, siendo los más utilizados el **riesgo relativo**, el **odds ratio**, la **reducción absoluta de riesgo**, y el **número de pacientes que será necesario tratar**. Los dos primeros son medidas relativas y los dos últimos medidas absolutas. Aunque en [artículo anterior](#) ya se comentaron estos parámetros, debido a la gran difusión de uno de ellos –el número necesario de pacientes a tratar–, vamos a profundizar un poco más en su significado, ventajas e inconvenientes.

Usaremos un ejemplo para la exposición. En un estudio de pacientes hipertensos se pretende investigar la posible asociación entre la probabilidad de padecer un infarto y el grado de control de la hipertensión. Para ello se analiza una muestra aleatoria de 728 pacientes, observándose los siguientes resultados

	Mal control	Buen control	
Infarto	10	7	17
No Infarto	250	461	711
	260	468	728

	% Infartos
Mal control	3.8
Buen Control	1.5

Tabla 1

El **número de pacientes que es necesario tratar** (abreviado NNT) se define como el número de pacientes que se estima que es necesario tratar con el nuevo tratamiento, en lugar de con el tratamiento control, para prevenir un suceso. Fue introducido por Laupacis y cols. en 1988, *como una medida clínica útil de las consecuencias de un tratamiento*. Desde entonces su utilización se ha ido popularizando, sobre todo a raíz del auge de la corriente de medicina basada en la mejor evidencia. Lo podemos comprobar de manera informal efectuando una búsqueda online por años en una revista general como [BMJ](#) de los artículos que contienen en el texto la frase "number needed to treat":

The screenshot shows a search engine interface with the following elements:

- Specify Authors, Keywords:** A section with input fields for "Author" (twice) and "Title". Below these are radio buttons for "Abstract | Title" and "Text | Abstract | Title".
- Limit Results:** A section with a "From" dropdown set to "Jan" and "2000", a "through" dropdown set to "Dec" and "2004", and an "Include:" section with radio buttons for "Articles only" and "Articles and abstracts only".
- Search Criteria:** The search terms are "number needed to treat".
- Options:** There are radio buttons for "any", "all", and "phrase" for both the search terms and the search criteria.

y obtenemos los siguientes y elocuentes resultados:

	Nº de artículos total
<b>1994</b>	1
<b>1995–1996</b>	24
<b>1997–1998</b>	54
<b>1999–2000</b>	56
<b>2001–2002</b>	63
<b>2003– Marzo 2004</b>	64

Tabla 2

El énfasis que se viene haciendo en la importancia de incorporar de forma crítica los resultados de la investigación científica en la práctica médica habitual, conlleva la necesidad de expresar los resultados de forma clara, inteligible y fácilmente trasladable a las situaciones que el médico encuentra en su actividad diaria; lo cual no siempre resulta sencillo, pues en primer lugar, tal y como se ha evidenciado en [numerosas ocasiones](#) el razonamiento probabilístico no entra por desgracia en el acervo cultural medio, no solo de los pacientes, sino tampoco del personal médico y de enfermería. Por otro lado, también se ha comprobado que la forma en la que se comunican los resultados influye decisivamente en la utilización que de éstos se hace, y en el impacto que producen en la práctica clínica. Los partidarios del NNT destacan de él precisamente esas cualidades: es fácil de calcular, fácil de comprender y tiene la ventaja de que no solo proporciona información estadística, sino también información clínica.

Sin embargo vamos a ver que las cosas no son tan claras como parece, y es posible que una vez más el amplio uso de este parámetro se deba más a una nueva moda, o a la pereza intelectual que nos lleva a imitar mecánicamente lo que vemos difundirse en las publicaciones.

Para el cálculo del NNT se utiliza como base la **reducción absoluta de riesgo**. En la [tabla 1](#) el riesgo de infarto en el grupo de mal control es del 3.8 % frente al 1.5 % en el grupo de buen control, y por lo tanto la reducción absoluta de riesgo (*RAR*) es la diferencia entre ambos, es decir 2.3. El NNT se calcula como

$$\text{Número Necesario a Tratar} = \frac{1}{\text{reducción absoluta del riesgo}}$$

Cuando la reducción se expresa en %, la fórmula se modifica como 100/*RAR*. En nuestro ejemplo

$$\text{NNT} = \frac{1}{\left(\frac{10}{260} - \frac{7}{468}\right)} = 42.5$$

Uno de los primeros "pequeños" inconvenientes que se presentan con el NNT es que puesto que hablamos de número de pacientes, parece lógico redondear a números enteros (de pacientes), y así se hace habitualmente. En nuestro ejemplo el valor del NNT, si redondeamos, es de 43 pacientes. Cuando la cifra obtenida es pequeña el efecto del redondeo puede ser importante, por ejemplo 3.5 pacientes pasa a ser 4. El argumento para redondear es que es más fácil de recordar una cifra entera. Mi opinión es que en los números bajos debiera mantenerse un decimal.

La idea de que hay que facilitar la asimilación de los datos puede conducir a la decisión errónea de exponer únicamente el valor del NNT obtenido, siendo éste, como cualquier otro parámetro estadístico, un resultado sujeto a incertidumbre, que por lo tanto debe ser expresamente cuantificada mediante el correspondiente intervalo de confianza.

Para calcular el **intervalo de confianza del NNT** podemos utilizar los valores del propio intervalo de confianza del RAR, que para muestras grandes se aproxima mediante la siguiente fórmula del error estándar para una diferencia de proporciones, basada en la distribución normal, aunque los programas de ordenador utilizarán habitualmente métodos más precisos (como el de Wilson):

$$Err. est. RAR = \sqrt{\frac{p1 \cdot (1-p1)}{n1} + \frac{p2 \cdot (1-p2)}{n2}}$$

El intervalo de confianza del 95% del RAR se calcula como  $RAR \pm 1.96 \text{ Err. est. (RAR)}$ .

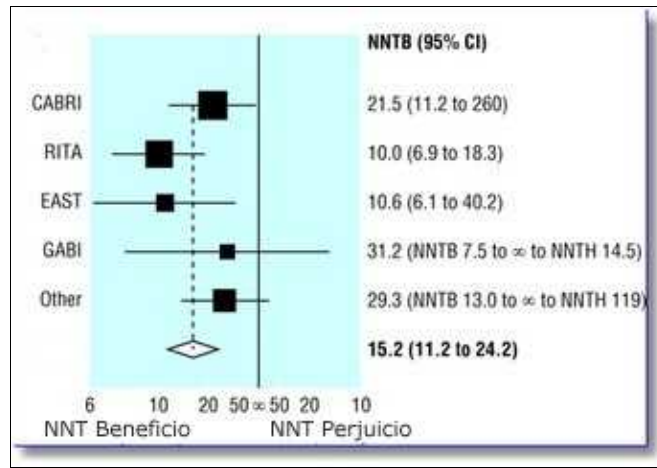
En nuestro ejemplo el error estándar expresado en porcentaje es de 1.31, con lo que el intervalo de confianza para el RAR va de -0.22 % a 4.92 %.

Si calculamos  $-100/0.22$  y  $100/4.92$  obtenemos el intervalo para el NNT, es decir -454.5 a 20.3. Y tenemos la primera sorpresa: un valor negativo. En cuanto nos paramos a pensar, vemos que la interpretación es muy simple. En este ejemplo el resultado no es estadísticamente significativo y por tanto el intervalo de confianza para la reducción de riesgo incluye el 0, va desde un valor negativo (ineficacia) a un valor positivo (eficacia). Cuando pasamos al NNT hablamos de número de pacientes a tratar para obtener beneficio y por el otro lado hablaríamos del número de pacientes necesario para perjudicar (se suele abreviar NNH, del inglés "harm", dañar). Entonces el intervalo de confianza es de  $NNT=20.3$  a  $NNH=455.5$ .

La segunda sorpresa es que el intervalo de confianza parece no incluir entonces el valor del NNT obtenido en nuestro estudio que era 42.5. Se debe a que en realidad la estimación nos dice que el número de pacientes que en promedio necesitaremos tratar puede ser tan bajo como 20 o crecer hasta infinito (no eficacia) y por lo tanto pasa por el valor obtenido de 43, o incluso perjudicar (preferible el grupo control).

Un poquito confuso. Mi opinión es que cuando no hay un efecto claro del tratamiento el concepto del número de pacientes a tratar para obtener beneficio no tiene sentido, y eso es lo que la matemática nos refleja. Y entonces ¿qué razón hay para no usar simplemente la reducción de riesgo cuyo IC va desde -0.22 % (mejor el control) a 4.92 % (mejor el tratamiento)?

En el artículo de [Altman](#), autor siempre muy didáctico, se explica con más detalle este aspecto de los intervalos de confianza para el NNT, y cómo presentar los resultados, en concreto cuando se están comparando varios estudios (meta-análisis), donde propone la siguiente gráfica:



Vemos que en el lado de la izquierda tenemos el número de pacientes para obtener beneficio y nos desplazamos hacia la derecha pasando por el punto de un número infinito de pacientes <sup>∞</sup> hasta la zona de perjuicio.

Cambiando de asunto ¿Por qué usar un índice de asociación absoluto como los anteriormente comentados (RAR o NNT) o, por el contrario emplear un índice relativo? Vamos a buscar respuestas a esa pregunta. Los índices relativos que se emplean habitualmente son el riesgo relativo y el odds ratio, ya [comentados anteriormente](#). La principal ventaja de los indicadores absolutos frente a los relativos radica en que aquellos nos dan una medida del impacto real del tratamiento. En el artículo de [Cook y Sacket](#) se presenta la siguiente tabla:

Hipertensión		ACVA en 5 años		Riesgo relativo Pc/Pt	Red.Abs.riesgo RAR	NNT
		Grupo control	Grupo tratamiento			
Moderada (TAS <=115)	Proporción sucesos	0.20	0.12	1.67	0.08	13
	Nº total pacientes	16778	16898			
Leve (TAS <=110)	Proporción sucesos	0.015	0.009	1.67	0.006	167
	Nº total pacientes	15165	15238			

Tabla 3

donde vemos que aunque el riesgo relativo 1.67 es el mismo en los dos grupos (hipertensión moderada y leve), sin embargo el impacto del tratamiento es mayor en el grupo de hipertensión moderada, donde por término medio se espera prevenir un evento tratando 13 pacientes, mientras que en el grupo de hipertensión leve es necesario tratar a 167. También podíamos haberlo expresado indicando que en el primer grupo el riesgo se reduce en un 8 % y en el segundo la reducción no llega al 1 % (0.6).

¿Por qué entonces se utiliza tanto el riesgo relativo y sobre todo, por qué se utiliza el odds ratio que es todavía más difícil de entender? Fundamentalmente por razones matemáticas, en concreto el odds ratio se deriva directamente del modelo logístico, y sobre todo cuando se van a combinar resultados de diferentes estudios,

por ejemplo en meta-análisis, ya que se ha evidenciado que los índices relativos tienden a ser más homogéneos entre diferentes ensayos del mismo tratamiento y están menos influenciados por el riesgo basal. Hay incluso [fuertes críticas](#) a la utilización indiscriminada del NNT obtenido a partir de meta-análisis.

### Cálculo del NNT en estudios de supervivencia

Cuando el dato de interés es el tiempo hasta que ocurre un suceso, como en un análisis de supervivencia, no podemos hablar de un único número de pacientes a tratar, sino que éste se puede calcular para cualquier instante de tiempo desde el inicio del tratamiento. El número de pacientes a tratar en este caso corresponde al número de pacientes que es necesario tratar para que un paciente adicional sobreviva hasta el tiempo para el que se calcula.

Si se está comparando la supervivencia en dos grupos, tratamiento y control, se habrá estimado la supervivencia en cada uno de ellos mediante el [método de Kaplan–Meier](#). Si deseamos calcular el NNT para un instante de tiempo a partir del comienzo del tratamiento, y el valor de la curva de supervivencia en ese instante es  $S_T$  y  $S_C$  para cada uno de los grupos, siendo  $S_T > S_C$  si el tratamiento es efectivo, la reducción absoluta de riesgo RAR vendrá dada como la diferencia entre ambos valores, por lo que el  $NNT = 1/(S_T - S_C)$ .

Conocidos los errores estándar de ambas curvas de supervivencia en ese punto, se puede calcular el error estándar de RAR como:

$$es(RAR) = \sqrt{es(S_T)^2 + es(S_C)^2}$$

donde es indica error estándar. Al igual que se explicó anteriormente, se utilizará este error estándar para calcular el intervalo de confianza para el RAR, y luego invirtiendo esos valores se obtiene el intervalo de confianza para el NNT.

Cuando las curvas de supervivencia se estiman mediante un [modelo de Cox](#), con el fin de tener en cuenta la influencia de otras variables, se puede estimar el NNT para un instante de tiempo dado, para un perfil de valores determinado de las otras variables que intervienen en el modelo. Si el coeficiente del modelo para la variable que especifica el grupo es  $b$ , siendo el valor de la variable correspondiente 1 cuando se trata del grupo de tratamiento, y 0 cuando se trata del grupo de control, la función de riesgo (hazard ratio) se estima como  $h = e^b$ . Si en un instante de tiempo la supervivencia en el grupo de control es  $S$ , en el grupo de tratamiento será  $S^h$ , por lo que el NNT vendrá dado por:

$$NNT = \frac{1}{S^h - S}$$

Para obtener el intervalo de confianza, se reemplaza en la ecuación anterior  $h$  por su intervalo de confianza que se calcula utilizando el error estándar del coeficiente  $b$  estimado para el modelo:

$$IC(h) = e^{b \pm 1.96es(b)}$$

Sin embargo hay que resaltar que este intervalo es optimista, ya que no se tiene en cuenta la imprecisión en la estimación de la supervivencia  $S$  en el grupo de control.

Cuando se desea calcular el NNT para análisis de supervivencia a partir de datos de la literatura no siempre se puede obtener de la información del artículo los valores anteriormente presentados para dicho cálculo. Es normal que se presente la curva de supervivencia en una gráfica, en la que de forma cuidadosa se puede leer los valores de dicha supervivencia y calcular a partir de ellos el NNT. Otra cosa es el asunto de disponer de los errores estándar, ya que éstos no suelen presentarse. Si se indica el intervalo de confianza del 95 % para la supervivencia, se puede calcular el error estándar dividiendo la amplitud del intervalo de confianza por 4:

$(S_{\text{SUB}} - S_{\text{INF}})/4$ .

Si tampoco disponemos del intervalo de confianza, lo que sí es habitual es que en el pie de la gráfica de supervivencia se especifique el número de pacientes que continúan vivos (conjunto de riesgo). Si llamamos  $n_t$  al número de pacientes que continúan vivos en el grupo de tratamiento, y  $n_c$  a los del grupo de control, podemos calcular de forma aproximada el error estándar de RAR como

$$es(\text{RAR}) = \sqrt{S_t^2(1 - S_t)/n_t + S_c^2(1 - S_c)/n_c}$$

(véase el [artículo sobre análisis de supervivencia](#) y el artículo de [Altman y Andersen](#) para más detalles)

## Epílogo

El NNT es un parámetro eficaz en términos de "potencia comunicadora" a la hora de presentar el impacto de un tratamiento o de una intervención, pero por ello mismo, debido a su capacidad de simplificación, hay que ser cautos a la hora de utilizarlo e interpretarlo, teniendo en cuenta que existen otras alternativas. Como cualquier resultado estadístico debe siempre venir acompañado de su intervalo de confianza, aunque éste resulta un tanto confuso si el tratamiento no es significativamente más eficaz, motivo por el cual en esos casos mis preferencias personales se inclinan por utilizar otros índices, como puede ser simplemente la reducción absoluta de riesgo.

El NNT es un parámetro previsto inicialmente para resumir resultados de tipo binario, por lo que el ansia de utilizarlo puede llevarnos a convertir un resultado numérico continuo en binario, dividiendo el resultado en dos categorías mediante un punto de corte, mejor o peor seleccionado. Siempre que se estratifica una variable continua se pierde información y se introduce sesgo, por lo que en general, aunque es una práctica bastante extendida por su simplicidad, se desaconseja su utilización, máxime cuando existen [métodos](#) para calcular el NNT para datos continuos (lógicamente más complejos).

Los NNT obtenidos a partir de un meta-análisis deben ser interpretados con cautela, debiendo siempre analizar el riesgo base en cada uno de los estudios, ya que si éstos son diferentes su utilidad es más que dudosa, y en esos casos es preferible utilizar un índice relativo como el odds ratio o el riesgo relativo.

Es preferible pensar en el NNT como un dato poblacional, que representa una comparación entre los grupos del esfuerzo clínico que hay que realizar, en promedio, para prevenir un suceso adverso, pero no se trata de un índice aplicable directamente de forma individual a un paciente con unas características concretas.

## Referencias

- ***Confidence intervals for the number needed to treat***  
D. G Altman  
BMJ, November 7, 1998; 317(7168): 1309 – 1312.  
[\[Full Text\]](#)
- ***Numbers needed to treat derived from meta-analyses---sometimes informative, usually misleading***  
L. Smeeth, A. Haines, and S. Ebrahim  
BMJ, June 5, 1999; 318(7197): 1548 – 1551.  
[\[Full Text\]](#)
- ***The number needed to treat: a clinically useful nomogram in its proper context***  
G. Chatellier, E. Zapletal, D. Lemaitre, J. Menard, and P. Degoulet  
BMJ, February 17, 1996; 312(7028): 426 – 429.  
[\[Full Text\]](#)
- ***The number needed to treat: a clinically useful measure of treatment effect***  
Richard J Cook, David L Sackett

BMJ 1995;310:452–454 (18 February)

[\[Full Text\]](#)

• Letters

[Risk measures expressed as frequencies may have a more rational response](#)

Peter Ayton

BMJ 1995;310:1269 (13 May)

• *Calculating the number needed to treat for trials where the outcome is time to an event*

D. G Altman and P. K. Andersen

BMJ, December 4, 1999; 319(7223): 1492 – 1495.

[\[Full Text\]](#)

• *Number needed to treat (NNT): estimation of a measure of clinical benefit*

Statistics in Medicine

Volume 20, Issue 24, Date: 30 December 2001, Pages: 3947–3962

S. D. Walter

[Abstract](#) | [References](#)



[Indice de artículos](#)

[Principio de la página](#)