

Number needed to treat: een nuttige maat voor de effectiviteit van een behandeling

P. N. VAN HARTEN

SAMENVATTING De concepten die ontwikkeld zijn in evidence-based medicine vormen naast klinische ervaring een belangrijke leidraad bij het nemen van klinische beslissingen. Een van die concepten is number needed to treat (NNT), het aantal patiënten dat behandeld moet worden om één extra ongunstige uitkomst te voorkomen of één extra gunstige uitkomst te bereiken. Het NNT is afhankelijk van het achtergrondrisico (het risico op de gebeurtenis bij aanvang): naarmate dit hoger is, zal het NNT van een gebleken effectieve interventie lager (dus gunstiger) zijn. Het achtergrondrisico heeft meestal geen invloed op de relatieve risicoreductie. NNT vormt samen met number needed to harm een getalsmatige argumentatie of een behandeling meer goed dan kwaad doet.

[TIJDSCHRIFT VOOR PSYCHIATRIE 50(2008)6, 337-343]

TREFWOORDEN achtergrondrisico, evidence-based medicine, number needed to harm, number needed to treat

Evidence-based medicine (EBM) is niet meer weg te denken uit de hedendaagse klinische praktijk (Offringa e.a. 2000; Sackett e.a. 1991 en 2000). De populariteit van EBM is mede te danken aan de grondleggers ervan, die vaak complexe epidemiologische begrippen vertaalden in praktische en didactische concepten (Sackett e.a. 1991 en 2000). Een van die concepten, het number needed to treat (NNT), werd in 1988 voor het eerst beschreven (Laupacis e.a. 1988). NNT sluit aan op een van de belangrijkste taken van de clinicus: een behandeling voorstellen die meer goed dan kwaad doet.

Clinici gebruiken woorden zoals 'vaak', 'soms' of 'meestal' om de effectiviteit van een behandeling te beschrijven. NNT is een numerieke waarde voor de effectiviteit van een behandeling. Een getal heeft het voordeel van meer eenduidigheid dan bovenstaande kwalificaties en is ook makkelijker overdraagbaar aan anderen. Echter, cijfers en berekeningen roepen bij veel behandelaren ook weer-

stand op. Daarom tracht ik in dit artikel om het concept NNT zodanig te bespreken dat een clinicus de praktische waarde ervan inziet en zal overwegen om het te gebruiken bij beslissingen rond een behandeling.

CASUS: THERAPIERESISTENTE PSYCHOSE

Bij een 21-jarige ziekenverzorger heeft zich een ernstige paranoïde psychose ontwikkeld waarbij zij zich achtervolgd voelt door mannen die haar zouden willen martelen en verkrachten. Ondanks behandeling met achtereenvolgens pimozide tot 4 mg, olanzapine tot 20 mg, risperidon tot 6 mg, en cisordinol tot 30 mg persisteert de psychose. Ook toevoegen van een benzodiazepine, een stemmingsstabilisator en een combinatie van twee antipsychotica geeft weinig verbetering. Zij lijdt zichtbaar en 's nachts komt zij soms doodsbang bij de verpleging. Na 6 maanden opname stelt de psy-

chiater voor om clozapine te proberen. Dit was al eerder aan bod gekomen, maar patiënte zag op tegen de wekelijkse bloedcontroles.

TE BEANTWOORDEN VRAAG

Het formuleren van een te beantwoorden vraag is de basis voor het oplossen van een klinisch probleem. Een nuttige benadering is het formuleren van een PICO-vraag, wat staat voor 'patiënt' (of probleem), 'interventie', 'comparison' (interventie waarmee vergeleken wordt) en uitkomst ('outcome'). De PICO-vraag is equivalent aan een onderzoeksvraag, en vergemakkelijkt het vinden van een antwoord in de wetenschappelijke literatuur (voorbeelden kunt u o.a. vinden op www.psychiatrienet.nl, klik op 'aandachtsgebieden' en dan op 'evidence-based medicine'; zie ook tabel 1).

De PICO-vraag kan voor deze casus als volgt worden geformuleerd: wat is de kans dat clozapine (interventie) vergeleken met een ander antipsychoticum (comparison) de psychose vermindert of doet verdwijnen (uitkomst), bij een patiënte die therapieresistent is (patiënt)? Therapieresistentie wordt gedefinieerd als persisteren van de psychose na twee behandelingen van ten minste zes weken met verschillende antipsychotica (niet clozapine).

BEREKENEN VAN HET NNT

NNT is het aantal patiënten dat behandeld moet worden om één extra ongunstige uitkomst te voorkomen of één extra gunstige uitkomst (genezing of remissie) meer te bereiken dan met de referentiebehandeling verkregen zou zijn

(Offringa e.a. 2000).

Vrij vertaald naar deze casus kan men het NNT als volgt formuleren: hoeveel therapieresistente patiënten moeten ingesteld worden op clozapine om één extra patiënt psychosevrij te maken vergeleken met het geven van een ander antipsychoticum?

Voor het berekenen van het NNT en de beantwoording van de PICO-vraag gebruiken wij een gerandomiseerd gecontroleerd onderzoek (Kane e.a. 1988) en een systematische review (Wahlbeck e.a. 1999).

In het onderzoek van Kane e.a. (1988) hadden patiënten net als bij deze casus meerdere antipsychotica gehad voordat gestart werd met clozapine. Een verschil was dat in dit onderzoek alleen antipsychotica van de eerste generatie gebruikt werden en de gepresenteerde patiënte meerdere antipsychotica van de tweede generatie had gekregen. In dit onderzoek vergeleek men de effectiviteit van clozapine met die van chloorpromazine: in de clozapinegroep reageerde 30% en in de chloorpromazinegroep 4% (Kane e.a. 1988). Het is gebruikelijk bij het vergelijken van risico's om uit te gaan van de slechte uitkomst, dus het percentage dat niet reageert. Het percentage dat niet reageert op clozapine of op chloorpromazine is dan respectievelijk 70 en 96.

Tabel 2 toont op welke wijze de absolute risicoreductie (ARR) en de relatieve risicoreductie (RRR) en het NNT berekend worden. NNT en ARR zijn reciproque waarden, ofwel NNT is 100 gedeeld door ARR. Zoals uit tabel 2 blijkt, is het NNT 4; er moeten dus 4 patiënten met een therapieresistente psychose behandeld worden met clozapine om bij één patiënt extra de psychose in remissie te bren-

| TABEL 1 | Nuttige websites met informatie over evidence-based medicine (EBM) |
|--|--|
| www.psychiatrienet.nl | Een onafhankelijke selectie van links van psychiaters voor psychiaters; onder 'aandachtsgebieden' staat de rubriek 'evidence-based medicine' |
| www.musc.edu/dc/icrebm/glossary.html | Site van Medical University of South Carolina, met uitleg van de basisbegrippen van EBM |
| www.cebmh.com/ | Met o.a. gratis programma om EBM-berekeningen te maken: CATmaker (van 'Critically appraised topic') en veel EBM-links |
| www.cochrane.nl/index.html | Site van het Nederlandse Cochrane-centrum, met verwijzing naar alle Cochrane-bronnen |
| www.jr2.ox.ac.uk/Bandolier/glossary.html | Britse site over gebruik van evidence in gezondheidszorg; aanbevolen zijn de 'extended essays' |

TABEL 2 Rekenvoorbeeld met de kans dat patiënten met een therapieresistente psychose herstellen met een behandeling met clozapine ten opzicht van behandeling met chloorpromazine (Kane e.a. 1988)

| Behandeling | Geen respons | Respons |
|---------------------------------|--------------|---------|
| Clozapine (in %) | 70 | 30 |
| Chloorpromazine (in %) | 96 | 4 |
| Absolute risicoreductie (in %) | 26 | |
| Relatieve risicoreductie (in %) | 27 | |
| Number needed to treat (n) | 4 | |

De kans op geen respons met clozapine is 70%.
 De kans op geen respons met chloorpromazine is 96%.
 Relatieve risicoreductie = $(96-70)/96 = 0,27 = 27\%$.
 Absolute risicoreductie = $96-70 = 26\%$.
 NNT clozapine = $100/ARR = 100/26 = 4$.

gen vergeleken met de behandeling met chloorpromazine. De getallen van de systematische review geven een vergelijkbaar NNT (Wahlbeck e.a. 1999).

Een ander voorbeeld: in een gerandomiseerd gecontroleerd onderzoek van patiënten met een depressie werd naast een cognitieve therapie een antidepressivum X of een placebo toegediend. Met cognitieve therapie plus het antidepressivum X herstelde 70% en met cognitieve therapie plus placebo herstelde 60%. Het NNT is dan 10 (berekening: $ARR = \text{aantal dat niet reageert op placebo} - \text{aantal dat niet reageert op antidepressivum X} = 40\% - 30\% = 10\%$, en $NNT = 100/ARR = 100/10 = 10$). Ofwel, van elke 10 patiënten die behalve met de cognitieve therapie behandeld worden met antidepressivum X zal 1 patiënt profijt hebben van dit antidepressivum. Men kan ook redeneren dat 1 op de 10 patiënten baat heeft bij middel X. Immers, van de 10 patiënten die behandeld werden met antidepressivum X hadden 9 die behandeling niet nodig, omdat 6 al herstelden met placebo en 3 niet herstelden ondanks gebruik van dit antidepressivum.

Een vergelijkbare maat is het *number needed to harm* (NNH): het aantal patiënten dat behandeld moet worden om één extra ongewenste uitkomst (een schadelijke bijwerking of overlijden) te krijgen vergeleken met de kans op die uitkomst in de referentiebehandeling. De berekening is analoog aan het berekenen van het NNT. Soms neemt men als uitkomstmaat 'kwaliteit van leven', waarmee

men de voor- en de nadelen van een behandeling integreert (Jones e.a. 2006).

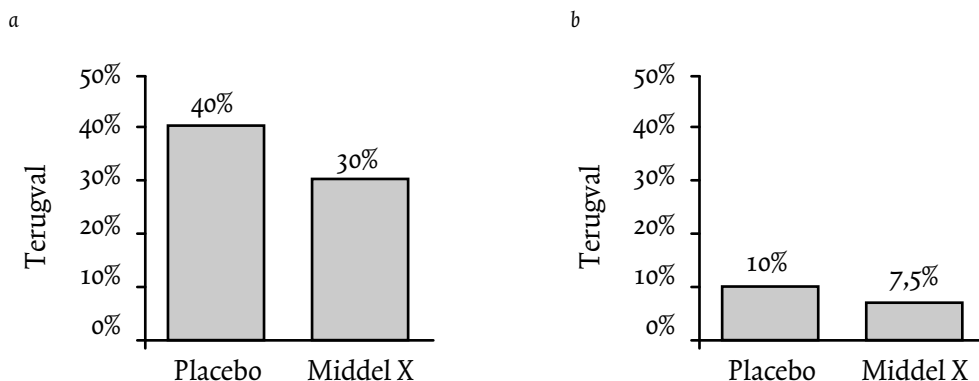
RELATIEVE VERSUS ABSOLUTE RISICOREDUCTIE

De effectiviteit van een behandeling kan op verschillende manieren beschreven worden. Commerciële instellingen zullen veelal de resultaten zo rooskleurig mogelijk weer willen geven en maken graag gebruik van de RRR. Om dit te verhelderen een voorbeeld: stel dat een ziekte slechts bij 1 op 1000 mensen voorkomt en dat het geven van medicijn X aan al die 1000 mensen de kans op de ziekte terugbrengt naar 1 op de 10.000. De RRR is dan maar liefst 90% (berekening: $(0,001-0,0001)/0,001 = 0,9$ is 90%). De ARR is echter miniem, slechts 0,09% en het NNT wordt dan 1111 (berekening: $100/ARR = 100/0,09 = 1111$). Het is uiteraard veel aantrekkelijker om aan te geven dat de kans op de ziekte met 90% afneemt dan dat er 1111 patiënten behandeld moeten worden om één slechte uitkomst minder te krijgen.

Artsen hebben echter belang bij een maat die objectief weergeeft hoe effectief de behandeling is ten opzichte van geen behandeling of een andere behandeling (Pinson & Gray 2003).

NNT is daar geschikt voor. Van groot belang voor de waarde van het NNT is het achtergrond-risico ofwel de prevalentie van de uitkomst (bijvoorbeeld een ziekte) in de populatie.

FIGUUR 1 Voorbeeld van hypothetisch onderzoek om invloed van achtergrondrisico te illustreren: kans op terugval met placebo versus middel X. In figuur 1a is de kans op terugval veel groter dan in figuur 1b, met andere woorden: in figuur 1a is het achtergrondrisico groter dan in figuur 1b (Baratt e.a. 2004)



Voor (a) geldt: kans op terugval met placebo is 40%; de kans op terugval met middel X is 30%. De absolute risicoreductie (ARR) is het verschil tussen beide: $40 - 30 = 10\%$.

De relatieve risicoreductie is het verschil in de kans op terugval in relatieve termen. Hier is de kans op terugval in de behandelde groep 25% minder dan in de placebogroep (10% verschil als proportie van de controlegroep = $10/40 = 25\%$ minder).

Het number needed to treat (NNT) is de reciproque waarde van de ARR, dus: $100/10 = 10$.

Voor (b) geldt: de kans op terugval met placebo is 10% en de kans op terugval met middel X is 7,5%. De absolute risicoreductie is dan $10 - 7,5 = 2,5\%$.

De relatieve risicoreductie is $2,5/10 = 0,25 = 25\%$. Het NNT is $100/ARR = 100/2,5 = 40$.

ACHTERGRONDRISICO

Het achtergrondrisico is de kans op de uitkomst in de controlegroep. Het NNT is gevoelig voor het achtergrondrisico, maar dit aspect van het NNT wordt onvoldoende onderkend (Offringa e.a. 2000; Sackett e.a. 1991 en 2000).

Een voorbeeld verduidelijkt veel: in figuur 1a zien wij dat 40% terugvalt onder behandeling met placebo en 30% onder behandeling met middel X. De ARR is 10% en het NNT is 10, ofwel: 10 patiënten moeten middel X krijgen om er één minder te laten terugvallen dan met placebo. Figuur 1b laat zien wat er gebeurt met de RRR, de ARR en het NNT als het achtergrondrisico daalt, ofwel de kans op terugval veel lager is: bij behandeling met hetzelfde middel X is de kans op terugval namelijk 7,5% en bij behandeling met placebo 10%. De RRR blijft gelijk (25%), maar de ARR wordt lager (nu

2,5%) en het NNT veel hoger. Nu moeten 40 patiënten middel X krijgen om er één minder te laten terugvallen dan met placebo. Algemeener gezegd: stel de kans op terugval in groep A is veel groter dan in groep B en een behandeling verlaagt de kans op terugval, dan zal de behandeling in groep A veel meer patiënten helpen dan in groep B. Daarom is het essentieel dat u nagaat of de kans op terugval in uw populatie overeenkomt met die in het onderzoek.

Dat het NNT en daarmee de effectiviteit van een behandeling afhangt van het achtergrondrisico lijkt bijzonder, maar ervaren klinici lijken hier intuïtief rekening mee te houden, zoals gesuggereerd wordt in het onderzoek van Snoek (1989). Hij vond dat bij het oplossen van diagnostische vragen ervaren klinici vaker dan hun onervaren vakgenoten gebruikmaakten van de volgende heuristiek: 'Als ik me bezighoud met de etiologie van een

symptoom, houd ik rekening met de frequentie van vóórkomen. De statistiek die ik hanteer, is voornamelijk gebaseerd op persoonlijke ervaringen'. Dit is min of meer te extrapoleren naar NNT, waarbij de waarde van het NNT afhangt van de frequentie van vóórkomen van de ziekte (of een andere uitkomst), ofwel het achtergrondrisico.

Het NNT wordt steeds meer gebruikt als effectmaat in meta-analyses en gerandomiseerd gecontroleerd onderzoek. In de CONSORT-richtlijnen (van *Consolidated standards of reporting trials*) is het advies om resultaten van klinische trials middels NNT's weer te geven (Altman e.a. 2001). NNT is concreter en klinisch beter te begrijpen dan de ook vaak gebruikte maten effectgrootte en oddsratio (Citrome & Stroup 2006; Gray 2004). Beide maten zijn om te zetten in NNT (Furukawa 1999; Sackett e.a. 1991).

Voor de patiënt is de vraag 'verbetert de kwaliteit van leven door de behandeling?' van cruciaal belang. Denk bijvoorbeeld aan de preventieve behandeling met antipsychotica om een terugval te voorkomen bij een patiënt die meerdere keren psychotisch is geweest. De effectiviteit van deze interventie is zowel statistisch significant als klinisch relevant, maar vanuit het perspectief van de patiënt kan de kwaliteit van leven zo afnemen dat deze het risico van terugval preferereert boven het nemen van een medicijn met bijwerkingen (Wunderink e.a. 2007).

BEPERKINGEN VAN NNT

Er zijn enkele beperkingen aan het concept NNT die van invloed zijn op de interpretatie.

Ten eerste is het NNT gebaseerd op een dichotome uitkomstmaat die vaak afkomstig is van een continue maat. Een voorbeeld kan dit verduidelijken: de beoordeling of een patiënt wel of niet hersteld is van een ziekte kan gebaseerd zijn op een meetschaal voor die ziekte. Meestal gebruikt men dan een afkappunt om te bepalen of de patiënt hersteld is of niet, maar het maakt natuurlijk uit of de patiënt een waarde heeft vlakbij of heel ver van het afkappunt. Die informatie verdwijnt ech-

ter bij het dichotomiseren en dus ook in het NNT.

Ten tweede wordt bij de waarde van NNT veelal een 95%-betrouwbaarheidsinterval gegeven. Dit betrouwbaarheidsinterval geeft aan dat met 95% zekerheid het NNT in het berekende interval ligt. Hoe groter het interval, hoe minder nauwkeurig het NNT bekend is. Als de effectiviteit van de ene behandeling niet significant verschilt van die van een andere, kan het berekenen van het 95%-betrouwbaarheidsinterval problemen geven. Zo kan de ondergrens van het betrouwbaarheidsinterval een negatieve waarde bevatten, terwijl het NNT alleen een positieve waarde kan hebben (Altman 1998).

Ten derde is onderzocht hoe patiënten het concept NNT ervaren. Er werd nagegaan of de hoogte van het NNT invloed had op de beslissing om een medicijn te nemen dat preventief werkte tegen hartinfarct en weinig bijwerkingen had. Bij een toename van NNT naar steeds hogere waarden (van 10 tot uiteindelijk 400) nam de bereidheid het medicijn te nemen wel significant af, maar die afname was gering. De interpretatie van de auteurs was dat patiënten het concept NNT niet goed begrijpen en er de voorkeur aan geven om de effectiviteit van een behandeling uitgelegd te krijgen in percentages verschil (Kristiansen e.a. 2002; Wen e.a. 2005). Deze interpretatie wordt echter ook betwijfeld (Akl 2004).

Ten vierde spelen methodologische vragen mee betreffende het gerandomiseerd gecontroleerd onderzoek waar het NNT uit gedestilleerd is. Zo is onder meer een vraag: komen de geïncludeerde patiënten in de gerandomiseerde gecontroleerde trial (RCT) overeen met de patiënten die bij de clinicus in behandeling zijn? Veelal zijn er bij RCT's veel uitsluitingscriteria en in de dagelijkse praktijk geldt dat vaak niet. Ook moet men zich afvragen: is de effectiviteit van de behandeling in de setting waarin de patiënt behandeld wordt even groot als in de setting van de RCT? Daar voert immers vaak speciaal getraind personeel de behandelingen uit. Tevens is de duur van het onderzoek van belang: deze is vaak korter dan de duur van een behandeling in de praktijk. Wordt na langere tijd

de behandeling meer of juist minder effectief? Dit heeft invloed op de werkelijke waarde van het NNT, maar dat vindt men niet in de relatief kortdurende RCT. Dat geldt ook voor bijwerkingen die pas na langere tijd optreden (bijvoorbeeld tardieve dyskinesie). Deze worden dan niet of onvoldoende gevonden gedurende het onderzoek en dus ook niet meegewogen.

CONCLUSIE

NNT is voor de clinicus een praktische maat die de effectiviteit van de behandeling weergeeft. De waarde van het NNT hangt samen met het achtergrondrisico, ofwel de prevalentie van de ziekte in de populatie die men onderzoekt. NNT en NNH vormen getalsmatige argumenten om te beoordelen of een behandeling meer goed dan kwaad doet. Naast deze exacte waarde spelen voor de interpretatie of een behandeling zinvol is ook persoonlijke of cultureel bepaalde opvattingen mee.

 Met dank aan dr. D.E. Tenback en P.R. Bakker, beiden psychiater bij de Symfora groep, voor commentaar op eerdere versies.

LITERATUUR

- Akl, E. (2004). Goodbye, number needed to treat? *Journal of Clinical Epidemiology*, 57, 219-220.
- Altman, D.G. (1998). Confidence intervals for the number needed to treat. *BMJ*, 317, 1309-1312.
- Altman, D.G., Schulz, K.F., Moher, D., e.a. CONSORT Group (Consolidated Standards of Reporting Trials) (2001). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Annals of Internal Medicine*, 134, 663-694.
- Baratt, A., Wyer, P.C., Hatala, R., e.a. (2004). Tips for learners of evidence-based medicine: 1. Relative risk reduction, absolute risk reduction and number needed to treat. *Canadian Medical Association Journal*, 171, 353-358.
- Citrome, L., & Stroup, T.S. (2006). Schizophrenia, Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) and number needed to treat: how can CATIE inform clinicians? *International Journal of Clinical Practice*, 60, 933-940.
- Furukawa, T.A. (1999). From effect size into number needed to treat. *Lancet*, 353, 1680.
- Gray, G.E. (2004). *Concise guide to evidence-based psychiatry*. Washington, DC/London: American Psychiatric Publishing.
- Jones, P.B., Barnes, T.R., Davies, L., e.a. (2006). Randomized controlled trial of the effect on quality of life of second- vs first-generation antipsychotic drugs in schizophrenia: Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS 1). *Archives of General Psychiatry*, 63, 1079-1087.
- Kane, J., Honigfeld, G., Singer, J., e.a. (1988). Clozapine for the treatment-resistant schizophrenic. A double-blind comparison with chlorpromazine. *Archives of General Psychiatry*, 45, 789-796.
- Kristiansen, I.S., Gyrd-Hansen, D., Nexøe, e.a. (2002). Number needed to treat: easily understood and intuitively meaningful? Theoretical considerations and a randomized trial. *Journal of Clinical Epidemiology*, 55, 888-892.
- Laupacis, A., Sackett, D.L., & Roberts, R.S. (1988). An assessment of clinically useful measures of the consequences of treatment. *New England Journal of Medicine*, 318, 1728-1733.
- Offringa, M., Assendelft, W.J.J., & Scholten, R.J.P.M. (2000). *Inleiding in evidence-based medicine*. Houten: Bohn Stafleu Van Loghum.
- Pinson, L., & Gray, G.E. (2003). Number needed to treat: an underused measure of treatment effect. *Psychopharmacology*, 54, 145-147.
- Sackett, D.L., Hayne, B.R., Guyatt, G.H. e.a. (1991). *Clinical epidemiology. A basic science for clinical medicine*. London: Little Brown.
- Sackett, D.L., Straus, S., Richardson, S., e.a. (2000). *Evidence-based medicine: how to practice and teach EBM*. London: Churchill Livingstone.
- Snoek, J.W. (1989). *Het denken van de neuroloog*. Groningen: Dijkhuizen van Zanten.
- Wahlbeck, K., Cheine, M., Essali, A., e.a. (1999). Evidence of clozapine's effectiveness in schizophrenia: a systematic review and meta-analysis of randomized trials. *The American Journal of Psychiatry*, 156, 990-999.
- Wen, L., Badgett, R., & Cornell, J. (2005). Number needed to treat: a descriptor for weighing therapeutic options. *American journal of health-system pharmacy*, 62, 2031-2036.
- Wunderink, L., Nienhuis, F.J., Sytema, S., e.a. (2007). Guided discontinuation versus maintenance treatment in remitted first-episode psychosis: relapse rates and functional outcome. *Journal of Clinical Psychiatry*, 68, 654-661.

AUTEUR

P.N. VAN HARTEN is hoofdredacteur van het Tijdschrift voor Psychiatrie en psychiater en A-opleider bij de Symfora groep, Amersfoort.

Correspondentieadres: dr. P.N. van Harten, Symfora groep, Postbus 3051, 3800 DB, Amersfoort.

E-mail: pn.van.harten@symfora.nl.

Geen strijdige belangen meegedeeld.

Het artikel werd voor publicatie geaccepteerd op 24-1-2008.

SUMMARY

Number needed to treat: a useful tool for estimating the effectiveness of a treatment – P.N. van Harten –

The concepts developed in evidence-based medicine, together with clinical experience, are the major principles on which clinical decisions are made. One of the concepts is number needed to treat (NNT), which means the number of patients who have to be treated either to prevent one additional unfavourable outcome or to achieve one additional favourable outcome. The NNT is dependent on the baseline risk (the risk of occurrence at baseline): the higher the baseline risk, the lower will be the NNT (which is more favourable). In contrast, the relative risk reduction is fairly independent of the baseline risk. NNT and number needed to harm provide numerical arguments for deciding whether a particular treatment does more good than harm.

[TIJDSCHRIFT VOOR PSYCHIATRIE 50(2008)6, 337-343]

KEY WORDS baseline risk, evidence based medicine, number needed to harm, number needed to treat