

Rationel FARMAKOTERAPI

11

Vurdering af behandlingseffekt ved kroniske sygdomstilstande



Er number-needed-to-treat det bedste effektmål?

Introduktion

En betydelig del af den praktiserende læges tid anvendes til at forebygge eller behandle kroniske sygdomme og tilstande som diabetes, hyperkolesterolemi, hypertension, osteoporose eller behandle uønskede konsekvenser heraf (hjerteinfarkt, apopleksi, øjenssygdom, nyresygdom, knoglebrud mv.). Det findes mange mulige interventioner, men lægen ønsker i alle tilfælde at vurdere effekten af interventionerne og informere patienten derom. Hensigten med denne artikel er at bidrage til en bedre forståelse af de kvantitative effekter af interventioner rettet mod kroniske tilstande, ikke at vurdere om behandlinger skal gives eller ikke.

Artiklen diskuterer måling af effekter på kliniske hændelser som infarkt, apopleksi eller død, ikke surrogatmål som blodtryks- eller kolesterolreduktion. Selv om det kan være interessant nok at kende sådanne effekter, er det først og fremmest effekterne på de relevante kliniske begivenheder (død, hjerteinfarkt, brud, etc.) og dermed livskvalitet, der bør afgøre behandlingsvalg. Dokumentation af sådanne effekter kræver randomiserede studier af store patientgrupper. Ved at sammenligne forekomsten af fx død eller hjerteinfarkt, kan man beregne effekterne af interventionen. Et godt eksempel på dette er et studie publiceret af Torp-Pedersen og medar-

bejdere, der randomiserede 6.676 patienter med nyopstået hjerteinfarkt til enten trandolapril eller placebo. Figur 1 viser, hvordan andelen af døde i de to behandlingsgrupper øges med tiden. Det synes åbenbart, at trandolapril bedrer patienternes overlevelse, men spørgsmålet er, hvordan vi bedst udtrykker denne forbedring.

Forskellige måder at udtrykke behandlingseffekt

I princippet kan behandlingseffekt måles »vertikalt« eller »horizontalt«. Ved den vertikale metode tager man udgangspunkt i andelen af døde på et givet tidspunkt, fx efter 5 års behandling. I Figur 1 er ca. 44% døde i trandolaprilgruppen og ca. 51% i placebogruppen. Vi vælger at kalde disse størrelser henholdsvis D_b (andelen døde i behandlingsgruppen) og D_k (andelen døde i kontrolgruppen). Effekten af behandlingen kan udtrykkes på forskellige måder:

1. Den absolutte risikoreduktion (ARR): $D_k - D_b$
2. Den relative risikoreduktion (RRR): $(D_k - D_b) / D_k$
3. Number-needed-to treat (NNT): $1 / \text{ARR}$

Desuden kan D_k og D_b benyttes til at beregne odds ratio ($OR = D_b (1 - D_k) / D_k (1 - D_b)$), som oftest (når D_k og D_b er små) vil være talmæssigt det samme som D_b / D_k . Både ARR, RRR og NNT er beregnet på grundlag af de samme

November 2003

**Vurdering af behandlingseffekt
ved kroniske sygdomstilstande
side 1**

Udgivet af
Institut for Rationel Farmakoterapi

Lægemiddelstyrelsen
Axel Heides Gade 1
DK-2300 København S

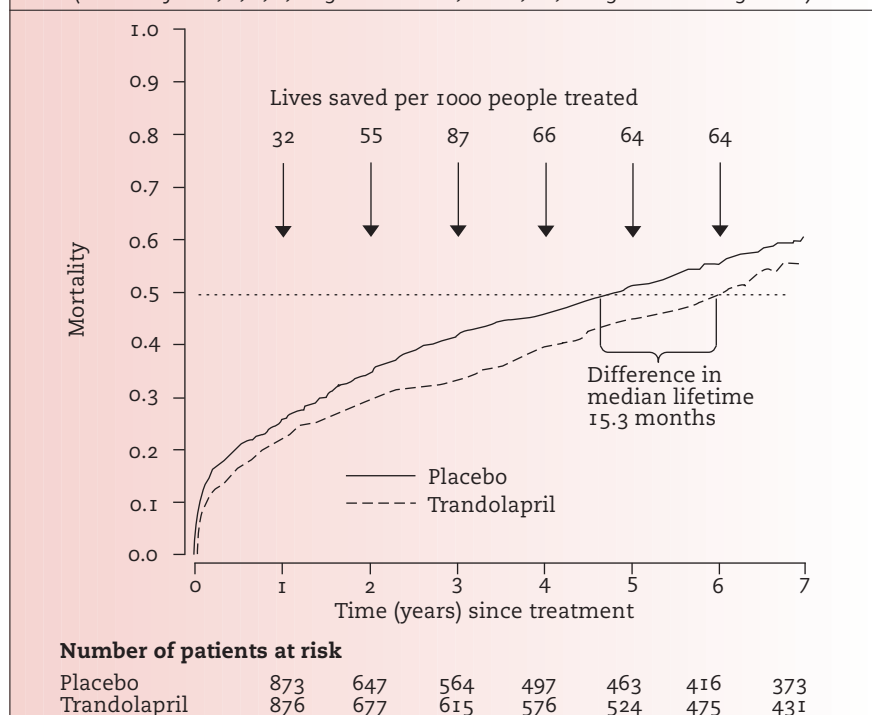
Tlf: 44 88 91 21
Man-tor 8.30-16.00
Fre 8.30-15.30
Fax: 44 88 91 22
E-mail: IRF@dkma.dk
<http://www.irf.dk>

Redaktion:
Institut for Rationel Farmakoterapi
Jens P. Kampmann (ansvarshavende)

Redaktionskomité:
Lars Bjerrum
Hanne Rolighed Christensen
Michael Dupont
Preben Holme Jørgensen

Sats og layout:
Lægeforeningens forlag
Tryk:
Scanprint, Viby J.
ISSN 1600-2555

Figur 1. Dødelighed ved trandolapril og placebo. Gengivet med tilladelse fra Elsevier Science (NNT er efter 1, 2, 3, 4, 5 og 6 år hhv. 31, 18, 11, 15, 16 og 16. Se i øvrigt tekst).



to andele (D_b og D_k), og det ene mål er således ikke mere korrekt end det andet, de udtrykker simpelthen forskellige aspekter af behandlingseffekten.

NNT er blevet et populært effektmål, fordi mange mener, det er nemt at forstå og intuitivt meningsfuldt. I trandolaprilstudiet angives »64 liv reddet ved behandling af 1.000 patienter«. Det betyder, at ARR er 6,4% efter 5 års opfølgning. Når ARR er 6,4%, er $NNT = 1/0,064 = 15,6 \sim 16$. Dette er populært blevet tolket, som man skal »behandle 16 patienter i 5 år for at redde et liv«. Men hvor mange har man reddet, hvis man behandler i 3 år eller i 6 år? Det kan man visuelt få et indtryk af ved at betragte den vertikale afstand mellem de to kurver. Jo større denne afstand er, jo større er ARR, og følgelig jo mindre er NNT. Ud fra Figur 1 kan man beregne, at NNT er 11 efter 3 års behandling og 16 efter 6 år. Det er således åbenbart, at både NNT, ARR og RRR er højst afhængige af, hvornår man måler effekten. I dette studie er NNT 31, 18, 11, 15, 16 og 16 efter henholdsvis 1, 2, 3, 4, 5 og 6 års behandling. De »vertikale« ef-

fektmål (ARR, RRR, NNT) er altså alle tidsafhængige i den forstand, at de udtrykker effekten målt på et enkelt tidspunkt. Trandolaprilstudiet viser, at selv om NNT er 16 efter 6 års behandling, er det ikke 16/2 efter den halve behandlingstid, men 11. Man kan ikke forudsige NNT ud fra, hvad man har observeret før eller senere i forløbet af behandlingen. Derfor er det problematisk at sammenligne NNT i forskellige studier med forskellig observationstid.

I denne artikel præsenterer vi bare NNT som en enkel værdi uden at angive sikkerhedsgrænser. Når man anvender det for at træffe beslutninger, bør NNT præsenteres med angivelse af sikkerhedsgrænser. Dette gøres sædvanligvis ved at tage den reciprokke værdi af sikkerhedsgrænserne for ARR. Hvis konfidensintervallet for ARR indeholder 0, bliver det mere kompliceret, fordi NNT da får to adskilte sæt sikkerhedsgrænser.

Det vigtigste »horizontale effektmål« er øgning (eller reduktion hvis interventionen har skadelig effekt) i median overlevelse. I trandolaprilstudiet er halvdelen af patienterne døde efter ca. 4,6 års op-

følgning. Den mediane levetid er altså 4,6 år for kontrolpatienterne, mens den er ca. 6 år i trandolaprilgruppen. Helt eksakt er den mediane overlevelsestid 15,3 måneder længere med trandolapril (Figur 1). Øgningen i middellevetid er et aggregeret effektmål, der kombinerer både vertikale og horisontale effekter. Denne øgning er repræsenteret ved arealet mellem to overlevelseskurver og kan beregnes matematisk ved integration eller empirisk ved at beregne den gennemsnitlige levetid i behandlings- og kontrolgruppen.

Ses effekterne over lang tid, bliver det klart, at strengt taget er ingen intervention livreddende, men kun livsforlængende. For den enkelte patient bliver spørgsmålet, hvor meget længere man lever, hvis man tager behandling. Dette spørgsmål besvares af overlevelseskurverne, som viser den gennemsnitlige forøgelse i middellevetid. Det er imidlertid ikke muligt at sige noget om, hvor meget den enkelte patient i studiet har forlænget livet, fordi små udskydelser af døden hos mange patienter vil give omtrent de samme overlevelseskurver som store udskydelser hos et fåtal.

Vi kan nu bedre forstå det vertikale effektmål NNT's begrænsning. At NNT er 16 efter 5 års behandling betyder ikke, at nogen har reddet livet. Det betyder, at nogen har forlænget livet en smule. Hvor meget kan man ikke sige uden at kende hele behandlingsforløbet. At NNT er 16, betyder heller ikke, at »de 15« ikke har haft nogen nytte af behandlingen. Hvor stor en andel, som har opnået en livsforlængelse, kan vi ikke vide, simpelthen fordi vi ikke kan forudsige, hvornår en patient vil dø hverken med eller uden behandling. Det er således ikke korrekt at sige, at »16 patienter skal behandles i 5 år for at undgå et dødsfald«. Det korrekte er, at »i gennemsnit skal 16 patienter behandles i 5 år for at observere ét dødsfald færre efter 5 år«.

I Figur 2 er vist et hypotetisk klinisk studie, hvor 10 patienter får behandling, mens 10 identiske patienter udgør en kontrolgruppe. Behandlingen udskyder døden et lille stykke tid hos alle 10, som får behandling. Imidlertid er der på et hvert tidspunkt en patient færre i live i kontrolgruppen. Derfor er NNT lig med 10, uanset hvornår man måler, selv om alle patienter har nytte af behandlingen.

Det er biologisk plausibelt, at interventioner mod kroniske tilstande som hyperlipidæmi forlænger livet hos størstedelen af de behandlede. Ræsonnementerne er de samme, når vi taler om hændelser som hoftebrud, hjerteinfarkt eller apopleksi, der ikke nødvendigvis er letale. Igen kan man tolke behandlingseffekten som udskydelse af disse uønskede hændelser hos en betydelig andel af patienterne.

Økonomisk evaluering og NNT

Det er en stor opgave at lave en fuldstændig økonomisk analyse af interventioner mod fx hyperlipidæmi eller osteoporose, og nogle har foreslået at anvende »poor man's cost-effectiveness analysis«: en analyse baseret på NNT. Hvis vi fx antager, at behandling med tran-dolapril koster 3.000 kroner per patient per år inklusive nødvendige blodprøver og lægebesøg, vil det koste omtrent 15.000 kroner at behandle en patient i 5 år. NNT efter 5 år er ca 16. Vi har altså behandlet 16 patienter for at »undgå« et dødsfald. Omkostningen per »undgået« dødsfald er altså $kr\ 15.000 \times 16 = 240.000$. Hvis vi foretager beregningerne efter 3 års behandling, er omkostningerne per patient 9.000 kroner, og NNT er da 11. Omkostningen per »undgået« dødsfald bliver nu 99.000 kroner. Omkostningen per undgået dødsfald er altså temmelig afhængig af, hvornår vi laver vores opgørelse. Konklusionen bliver altså, at NNT i almindelighed ikke er egnet som basis for økonomisk evaluering.

Hvilke effektmål er nemmest at forstå?

Der er lavet talrige studier, der viser, at valg af effektmål påvirker beslutningstageres behandlingsvalg. Det er vist, at såvel patienter som læger og politikere oftere foretrækker en behandling, når effekten udtrykkes som RRR end NNT eller ARR. Så længe der ikke findes en guldstandard for, hvad der er rigtigt og forkert, er det ikke muligt at sige, om det er RRR, som »overdriver« effekten, eller det er NNT, som »undervurderer« den. De er begge baseret på de samme to andele (D_k og D_b) og er begge sande mål for behandlingseffekt. I forhold til at informere om alle tre vertikale effektmål samtidig, bliver effekten »overvurderet«, når patienterne bare får RRR at vide, og »undervurderet«, om de bare oplyses om NNT.

Odense Risk Group har lavet studier af, hvordan læg og lærd tolker de forskellige effektmål. Disse undersøgelser viser entydigt, at lægfolk har svært ved at forstå NNT, samt at de er lige så positive over for en behandling, der har NNT på 200 eller 400, som til en med NNT på 10. Mange tror også, at »NNT på 16« betyder, at »15 ikke har nytte af behandlingen«. Hvis der er tale om en behandling uden tidsdimension (effekten ses umid-

FAKTA

Effekten af 5 års behandling af patienter med koronar hjertesygdom med simvastatin kan udtrykkes på følgende måder:

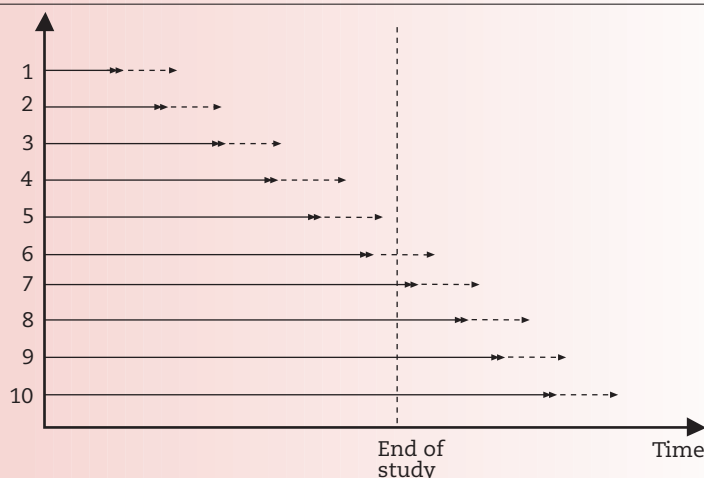
- RRR = 30%
- ARR = 3,3%
- NNT = 30

Gennemsnitlig livsforlængelse: 2-3 måneder

delbart), som fx DC-konvertering af ventrikelflimmer, er udsagnet naturligvis sandt. Når det gælder kroniske sygdomstilstande er det mere kompliceret. Vi mener således, at både læg og lærd har problemer med at forstå de vertikale effektmål. Ud fra det, som publiceres i den medicinske litteratur, mistænker vi, at få tænker på, at NNT kan variere meget efter, hvornår den bliver målt.

Hvad så med udskydelse som effektmål? Man skulle tro, at både læg og lærd forstår, hvad det betyder, at 5 års behandling med statiner i gennemsnit forlænger patienternes liv med 2-3 måneder. I den forbindelse skal man huske på, at vi ikke ved, om alle patienter får 2-3 måneder, eller om en større effekt er fordelt på en mindre andel. Noget sværere kan det være at forestille sig effekten af interventioner som udskydelse i tid af fx hjerteinfarkt eller hoftebrud, men i en dansk undersøgelse vedrørende

Figur 2. Klinisk randomiseret forsøg med sammenligning af kontrol- og behandlingsgruppe. Fuldt optrukken linje angiver levetid i kontrolgruppe. Den punkterede linje angiver levetidsforlængelse ved behandling. End of study angiver tidspunkt, hvor studiet gøres op.



Postbesørget blad 57179 (8245 ARC)

Adressemærkaten er udskrevet fra Lægeforeningens edb-register. Adresseændringer m.v. bedes meddelt til DADL, Registreringsafdelingen, Esplanaden 8 C, 1263 København K.

osteoporose kunne lægfolk godt skelne 1 måneds udskydelse af hoftebrud fra 2 års.

Hvilket effektmål er bedst?

Trods den store popularitet NNT har fået, findes ingen evidens for, at NNT bidrager til bedre beslutninger end andre effektmål. At beslutningstagerne er mere skeptiske til behandling, når effekten præsenteres som NNT end som RRR (Figur 3) beviser, at valg af effektmål har betydning, men ikke hvilket, der er bedst. Ingen kan i dag med sikkerhed sige, hvilke effektmål der bedst er i stand til at udtrykke effekten af interventioner mod kroniske sygdomstilstande. Det er centralt, at effektmålene let misforstås, specielt hvis man ikke tænker på tidsdimensionen. Effekten af en intervention er størst i grupper af patienter, der er mest udsatte for de uønskede kliniske konsekvenser. Det betyder, at behandlingseffekterne er størst i højrisikogrupper, uanset om man udtrykker effekten med vertikale eller horisontale effektmål (Tabel 1). Medicinsk behandling kan ikke helt fjerne patienternes risiko for at få hoftebrud eller hjerteinfarkt, men hvis behandlingen udskyder disse hændelser længe nok, når de måske at dø af andre årsager, inden de får infarkt eller hoftebrud.

Konklusion

Tidsdimensionen er essentiel ved vurdering af intervention på kroniske sygdomsprocesser. Mens de horisontale effektmål beskriver ag-

Tabel 1. Når basisrisikoen i kontrolgruppen er stor, har et større antal personer gavn af en behandling, end hvis basisrisikoen er lille.

Tilfælde	Kontrolgruppe	Behandlingsgruppe	RRR	ARR	NNT
1	30%	5%	83,3%	25%	4
2	84%	14%	83,3%	70%	2
3	10%	1,7%	83,3%	8,3%	13
4	95%	70%	26%	25%	4

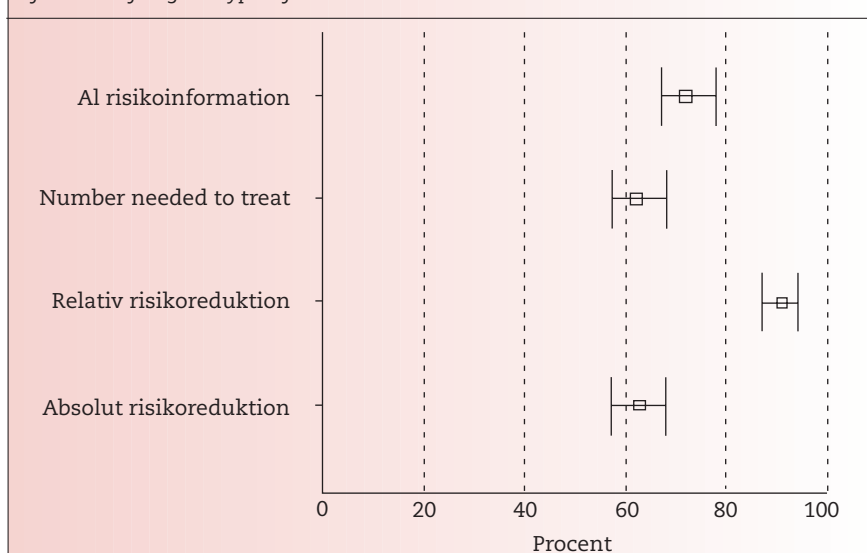
gregerede effekter over tid, mangler de vertikale effektmål (RRR, ARR, NNT) en tidsdimension. Disse effektmål kan variere meget over tid, og man kan groft fejlvurdere effekten ved at lægge vægt på et enkelt tidspunkt. Der er ingen evidens for, at NNT hjælper patienter eller læger til at træffe bedre beslutninger end andre effektmål. Derimod er det efterhånden godt dokumenteret, at såvel læg som lærd har svært ved at forstå NNT eller direkte misforstår betydning

af det. NNT bør som alle effektmål tolkes med varsomhed.

Ivar Sønbo Kristiansen
Palle Mark Christensen
Jørgen Nexø
Institut for Sundhedstjenesteforskning
Syddansk Universitet, Odense

Conflict of interest: Vore egne studier inden for dette område er finansieret af Hjertefonden, Forskningsrådet, Nycomed Pharma og PLU-fonden (Praktiserende Lægers Uddannelses- og udviklingsfond). Forfatterne har ikke haft tilknytning til producenten af trandolapril eller forskergruppen, som har offentliggjort trandolaprilstudiet, vi bruger som eksempel i artikelen.

Figur 3. Andel af læger, der helt sikkert eller formentlig vil anbefale lægemidlet som en funktion af angivet type information om risikoreduktion.



Kursuskataloget 2004

Forårets kurser kan nu ses på vores hjemmeside www.irf.dk.

Tilmelding til kurserne foregår udelukkende via www.irf.dk og er mulig fra mandag den 24. november 2003.