

$$\frac{\text{Effekt}}{\text{Bivirkninger}} \times \text{pris} = \text{Rationel Farmakoterapi}$$

Sammensatte endepunkter – Hvordan fortolkes de?

Af Thalia M. Blicher og Jens Peter Kampmann, IRF

I store undersøgelser er det almindeligt at anvende sammensatte endepunkter – ofte både 5-7 – om primære effekt/sikkerhedsparametre, frem for enkeltstående endepunkter.

Hvad vil du helst: Dø eller blive indlagt på hospital?

Et lægemiddels effekt og sikkerhed vurderes – både af godkendelsesmyndigheder og den enkelte læge – på baggrund af en ændring i præspecifiserede endepunkter eller variable som fx død, AMI, tid til hospitalisering etc. Disse hændelser skal optræde et antal gange, hvis de med statistisk sikkerhed skal kunne sige noget om en sammenkobling mellem behandling og hændelse. Er der tale om relativt sjældne hændelser, som død heldigvis er, kræver det som regel, at der indgår et stort antal patienter i undersøgelsen. Det koster både tid og penge.

10 hændelser på 100 patienter over for 5 på 100 er ikke signifikant, mens 100 hændelser på 1.000 over for 50 hændelser på 1.000 er signifikant. Samme eventhyppighed, men fastholdt over for et større antal patienter er sikrere end vist på kun få patienter.

Sammensatte endepunkter kombinerer flere endepunkter, som måles samtidigt. Dette medfører samlet set

en øget hændelsesfrekvens og dermed behov for færre antal patienter og/eller kortere opfølgningstid. Hvis man kombinerer hændelserne A og B, skal en af dem optræde for at blive registreret som en observation.

Fortalerne for denne metode hævder, at såfremt man vælger et acceptabelt sammensat endepunkt, får man et bedre helhedsbillede af sygdommen, mens omvendt modstandere spørger, hvem der skal afgøre, hvad der er acceptabelt. Undersøgelserne bliver sværere at sammenligne, da de næsten aldrig anvender helt ens endepunkter, og hvad gør man, når endepunkterne går hver sin vej. Hvor mange skal gå i samme retning? Kan et fald i fx ikke-letale myokardieinfarkter opvejes af en stigning i letale myokardieinfarkter – eller omvendt?

Eksempel: DREAM studiet (1) viste, at rosiglitazon på det sammensatte endepunkt signifikant reducerede risikoen for diabetes eller død. Det lyder jo meget godt, men ser man på de enkelte endepunkter, viser det sig, at rosiglitazon kun har vist effekt på udviklingen af diabetes, hvilket også var forventeligt. Der var ingen effekt på død.

De fleste patienter vil formentlig ikke tillægge diabetes og død samme betydning. Det er reduktion i fore-

komst af diabetes, der bærer signifikansen i det sammensatte endepunkt, og hvis læseren ikke er opmærksom på dette, kan studiet blive mistolket.

Baggrund

Freemantle et al (2) fandt i 9 store tidsskifter fra 1997 til 2001 167 randomiserede studier, hvor død indgik i den primære sammensatte effektparameter. Tilsammen var randomiseret ca. 300.000 patienter, og halvdelen af studierne undersøgte kardiovaskulær sygdom. De sammensatte endepunkter er altså et meget almindeligt problem, som alle, der udfører og læser randomiserede undersøgelser, må have en holdning til.

Ferreira-Gonzalez et al (3) fandt ved en systematisk gennemgang af 114 kardiovaskulære studier med sammensatte endepunkter, at de for patienten mindst vigtige endepunkter ofte har den højeste eventrate. Høje eventrater og større behandlingseffekt på mindre vigtige endepunkter kan være vildledende.

Hvad siger »loven«?

Inden for EU, USA og Japan foregår der et internationalt registreringsarbejde, der er udmøntet i de såkaldte ICH-regler (4) (*International Conference on Harmonisation*). I ICH

E9: »*Statistical Principles for Clinical Trials*« defineres et sammensat endepunkt som følger: »... hvis en enkelt primær parameter ikke kan vælges ud fra multiple hændelser associeret med studiets primære formål, kan man på baggrund af en prædefineret algoritme kombinere de multiple hændelser til én komposit variabel. P-værdien skal ikke justeres«. Dette er i modsætning til studier med multiple endepunkter, hvor p-værdien skal korrigeres afhængigt af antal endepunkter. Dette er altså ikke nødvendigt for de sammensatte endepunkter, men resultaterne kan være vanskelige at fortolke (5). Der står intet om, hvordan komponenterne i det sammensatte endepunkt skal behandles. FDA (6) (*food and drug administration*) i USA skriver i sine retningslinjer, at beskrivelsen af komponenterne i det sammensatte endepunkt kun bør være deskriptiv. Det vil sige, at komponenten ikke skal fremstilles med statistiske analyser, medmindre den er prædefineret som et endepunkt med en forud defineret hypotese og analyseplan.

Statistik

I sin simple form beregnes antal patienter ud fra forventningen om den samlede eventhyppighed for summen af alle delendepunkterne. Undersøgelsen får derfor styrke til at vise signifikans for det samlede endepunkt, men ikke for delendepunkterne, medmindre det er forudbestemt, og der er taget højde for multiplicitet. Det er uafklaret, hvor mange af delendepunkterne, der skal være signifikante for at give en overordnet signifikans. Skal det ikke være nogen som helst eller de fleste? Eller kun de allermest relevante? Man kan forestille sig, at det sammensatte endepunkt ikke er statistisk signifikant, kan man så overhovedet gå videre og analysere delendepunkterne? Og hvis det sammensatte endepunkt er statistisk signifikant, er delendepunkterne da prospektivt defineret med tilpas mange patienter til at sige noget endeligt? Ingen ved det.

Man kan gå så vidt, som Moyé (7) har gjort det, idet han »forlanger« at fordele p-værdien mellem det primære endepunkt og delendepunkterne, men det vil stort set medføre, at hele fordelingen ved det sammensatte endepunkt forsvinder. Samme forfatter introducerer løsninger til denne problemstilling, således at nogle af fordelene ved det sammensatte endepunkt bevares, men problemet er simpelthen ikke løst statistisk tilfredsstillende.

Diskussion

Hvis man ikke stringent kan behandle problematikken »sammensatte endepunkter«, er der i al fald nogle mere generelle retningslinjer, man kan tage hensyn til (8, 9, 10, 11, 12):

1. Hvor mange endepunkter indgår i det sammensatte endepunkt? Sjældent ses flere end seks. I analysen indgår alle endepunkterne med lige stor vægt og er alle relevante (kvalitativ homogenitet). Det vil sige, at alle events – både letale og ikke-letale – tillægges samme betydning. Patienterne derimod vil som regel hellere genindlægges på hospital end dø. Og hvordan med forskellige ikke-letale events? Er de valide surrogater? Validitet vil kræve, at der er en stærk sammenhæng mellem de ikke-letale endepunkter i den patofysiologiske kaskade af hændelser og efterfølgende død, for eksempel AMI og død. Denne sammenhæng eksisterer ikke for alle ikke-letale endepunkter: Ventrikulære ekstrasystoler hos post-AMI patienter øger risikoen for pludselig død. Man troede, at man ved at eliminere disse ekstrasystoler med klasse 1a antiarytmika kunne reducere mortaliteten. Resultatet blev dog lige omvendt, og behandlingen medførte øget mortalitet, da disse antiarytmika ikke hjælper på den underliggende strukturelle skade i hjertet.
2. Alle endepunkterne skal selvfølgelig være præspecificerede, lagt

klart frem og være tydeligt vurderbare. Man skal have mulighed for at se, om det er et enkelt eller få endepunkter, der dominerer det samlede endepunkt. Helt galt går det, hvis det mindst relevante endepunkt kvantitativ er det mest dominerende (kvantitativ heterogenitet). Der skal være nogenlunde samme relative risikoreduktioner for alle endepunkterne.

3. Der kan opstå »konkurrence« mellem endepunkterne. En forsøgsperson, der dør, kan ikke få et myokardieinfarkt eller blive genindlagt. Falder incidensen af AMI eller genindlæggelse, kan det skyldes, at incidensen af død stiger.
4. To endepunkter kan dække over samme eller næsten samme begivenhed. Kan man fx registrere både blodsukker og HbA_{1c} som delendepunkter?
5. Hvad sker der med patienterne, der får en event? En patient får fx 2 AMI og 1 apoplexi. Er det altid den først opståede event, der skal medtælles, eller kan andre kriterier gælde, fx alvorligheden af begivenheden?

Konklusion

Principper for anvendelsen af sammensatte endepunkter må herefter være:

1. Alle endepunkter skal være præspecificerede og klinisk relevante og skal fremgå af artiklen.
2. Endepunkterne må hverken være for forskellige eller for ens. De bør repræsentere forskellige aspekter af samme patofysiologiske proces og kunne korreleres til hinanden og til det letale endepunkt.
3. Endepunktets relative risikoreduktion skal være nogenlunde ens, og deres signifikansniveau bør være defineret fra starten.
4. Effekten af den behandling, der afprøves, skal være nogenlunde ens over for alle endepunkter, der indgår i det sammensatte endepunkt, og man skal være opmærksom på, hvilken eller hvilke endepunkter

der bærer »signifikansen« for det sammensatte endepunkt.

Memento

Husk på, at en undersøgelse med et sammensat endepunkt ikke nødvendigvis har vist signifikans for de enkelte delendepunkter; så man har fx ikke vist, at et lægemiddel reducerer letaliteten, hvis dette endepunkt i sig selv ikke er signifikant, men kun signifikant slået sammen med andre ikke-letale endepunkter. Hvis dette i skyndingen glemmes, kan undersøgelser fortolkes helt forkert. De bliver ikke nemmere at rapportere.

Referencer

1. The DREAM Trial Investigators: Gerstein HC, Yusuf S, Bosch J et al. Effect of rosiglitazone on the frequency of diabetes in patients with impaired glucose tolerance or impaired fasting glucose: a randomised controlled trial. *Lancet* 2006;368:1096-1105.
2. Freemantle N, Calvert M, Wood J, Eastaugh J and Griffin C. Composite outcomes in randomized trials: Greater precision but with greater uncertainty? *JAMA* 2003; 289(19):2554-59.
3. Ferreira-Gonzalez I, Busse JW et al. Problems with the use of composite endpoints in cardiovascular trials: systematic review of randomised controlled trials. *BMJ* 2007;334:786.
4. The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). www.ich.org
5. Schulz KF and Grimes DA. Multiplicity in randomised trials: endpoints and treatments. *Lancet* 2005; 365:1591-5.
6. US Food and Drug Administration. Guidance for Industry: Clinical studies section of labeling for human prescription drug and biological products - content and format.
7. Moyé LA. Multiple Comparisons in Clinical Trials. Springer, 2003; ISBN 038700727X, 9780387007274.
8. Lauer MS and Topol EJ. Clinical Trials - Multiple Treatments, Multiple End Points, and Multiple Lessons. *JAMA* 2003;289(19):2575-7.
9. Lubsen J and Kirwan B. Combined endpoints: can we use them? *Statistics in Medicine* 2002;21:2959-70.
10. Cannon CP. Clinical Perspectives on the use of Composite Endpoints. *Control Clin Trials* 1997;18:517-29.
11. Montori VM, Busse J, Permyer-Miranda G, Ferreira I and Guyatt G. How should clinicians interpret results reflecting the effect of an intervention on composite endpoints: should I dump this slump. *ACP Journal Club*. Nov/Dec 2005, vol. 143 Issue 3, pA8-A9.
12. Ferreria-Gonzalez I, Permyer-Miranda G, Busse JW, Bryant DM, Montori VM, Alonso-Coello P, Walter SD and Guyatt GH. Methodologic discussions for using and interpreting composite endpoints are limited, but still identify major concerns. *J Clin Epidemiol* 2007; 60:651-7.

Hvornår skal man starte behandling med statin mod højt kolesterol?

Af Niels C. Heebøll-Nielsen, IRF.

Talrige undersøgelser har vist, at det er muligt at reducere morbiditet og mortalitet af hjerte-karsygdom ved at behandle personer med øget risiko med statin (1).

Det er også kendt, at incidensen af hjerte-karsygdom (CVD) stiger med stigende S-kolesterol, men der er intet knæk på kurven, hvorfor det ikke er muligt at fastsætte en grænse for, hvornår det er fornuftigt at starte behandling af højt kolesterol med statin.

Litteraturen om værdien af primær profylakse med statin er sparsom. Der findes ingen store randomiserede undersøgelser, som har demonstreret signifikant effekt af statinbehandling af personer uden tegn på CVD. En ny metaanalyse om primær profylakse ved behandling med statiner er publiceret (2). Analysen omfatter 19 un-

dersøgelser, som alle viser insignifikant effekt af primær profylakse, men når undersøgelserne opgøres samlet, findes en relativ risiko på 0,93 med 95 % sikkerhedsgrænser 0,87 til 0,99. Der er blot det problem, at undersøgelserne, som indgår i analysen, handler om patienter, der faktisk er syge (diabetikere og personer med perifer arteriel insufficiens) eller i øvrigt har stærkt øget risiko for hjerte-karsygdom (plaque i carotiderne, nyre- og pancreastransplantation mv). Analysen kan ikke tages som udtryk for effekten af statinbehandling af raske personer, altså egentlig primær profylakse.

Måling af S-kolesterol alene er en dårlig prædikator for CVD. Østerbroundersøgelsen viste, at den gennemsnitlige kolesterolværdi for 50-årige

mænd var 6,14 og for kvinder 6,05 mmol/l. Ikke desto mindre angives behandlingsgrænsen for hyperkolesterolæmi flere steder som >5,0 mmol/l. Langt hovedparten af den danske befolkning er således patienter og skal behandles!

Ifølge DSAM's vejledning, skal man beregne den enkelte persons risiko for at udvikle CVD, før man beslutter sig til at behandle hyperkolesterolæmi (3). I denne beregning indgår flere forhold end S-kolesterol: Køn, alder, rygevaner, blodtryk, familiær disposition for CVD. Dertil oplysninger om væsentlige risikofaktorer som diabetes, perifer arteriel insufficiens, cerebrovaskulær sygdom og familiær hyperkolesterolæmi, som skal behandles særskilt med sekundærprofylakse.

En risikovurdering kan beregnes med SCORE-systemet (Systematic CORonary Risk Evaluation), som anført i DSAM-vejledningen.

Personer uden tegn på CVD og diabetes med en absolut 10-års risiko for at dø af CVD under 5 % bør almindeligvis ikke behandles med statin uanset kolesterolværdien.

Er risikoen over 5 % vil man i første omgang afvente resultatet af råd om livstilsændringer, før man giver statin.

Ved risiko over 10 % vil det næsten altid være nødvendigt at give statin.

Konklusion

En isoleret måling af kolesterol giver ikke indikation for behandling med statin.

En behandlingsgrænse på 5,0 mmol/l giver ikke mening, med mindre der er tale om manifest kardiovaskulær sygdom eller diabetes.

Referencer

1. Baigent C, Keech A, Kearney PM et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet*. 2005;366:1267-78.
2. Mills EJ, Rachlis B, Wu P et al. Primary prevention of cardiovascular mortality and events with statin treatments: a network meta-analysis involving more than 65,000 patients. *J Am Coll Cardiol*. 2008;52:1769-81.
3. Dansk Selskab for Almen Medicin 2007. Forebyggelse af iskæmisk hjerte-kar-sygdom i almen praksis.

Voldsom prisstigning på glimepirid

Ved de nyeste prisændringer observeres en voldsom prisstigning på glimepirid 3 mg tabletter. Glimepirid er i øjeblikket 1. valg på mange af regionernes basislister. Fra sidste takstperiode ses en 10-dobling af prisen – fra kr. 0,31 til kr. 3,61 per tablet. Antager man, at prisen fastholdes på dette niveau betyder det en yderligere udgift per patient på kr. 1.200 per år. Ca. 45.000 patienter anvender glimepirid i øjeblikket, men i forskellige doser. I 2008 blev der solgt 30.200 pakninger af glimepirid 3 mg. Det betyder med den aktuelle prisstigning en årlig merudgift på knap kr. 9 mio.

Tabletter med 1, 2 og 4 mg koster fortsat under 50 øre per tablet, hvilket kan give anledning til overvejelser om skift. Diabetes er en sygdom, som ofte kræver polyfarmaci, så forsøg med at skifte fra den dyre 3 mg til en kombination af tabletter af hhv. 1 og 2 mg må generelt frårådes af hensyn til patientens compliance.

Til gengæld kunne prisstigningen være en god anledning til at revurdere patientens dosis. Sulfonylurinstofferne har en meget flad dosisresponskurve, som bevirker, at man opnår begrænset mereffekt – om nogen – ved at øge dosis. Det anbefales

fålede dosisinterval for glimepirid er 1-4 mg, men der er sjældent grund til at anvende mere end 1-2 mg daglig.

IRF's formål med denne artikel er dog primært at skærpe opmærksomheden på disse – til tider – kritisable prisændringer, som følger i kølvandet på rationelle tiltag som fx regionernes basislister. Heldigvis falder prisen ofte igen, som det fx var tilfældet i det tidligere priscirkus på statiner.

Er artiklerne for svære at læse

En praktiserende kollega har påtalt, at teksten i Rationel Farmakoterapi »var meget tung og propet med fagterminologi, som jeg ikke forstod«. Det er artiklen i aprilnummeret om behandling af acne vulgaris, han henviste til. Han havde talt med flere kolleger, som delte hans opfattelse. Han bad om en »Pixie-udgave« af artiklen.

Rationel Farmakoterapi er et fagtidsskrift, som vi håber, at mange læger har glæde af at læse. Hvis sproget opfattes som en hindring, har vi brug for at få det at vide. Send os en mail på irf@dkma.dk, hvis du har kommentarer til problemet.

