

Vetenskapligt Förhållningssätt

Välkommen till utbildning i kritisk litteraturläsning

Syftet är att du ska lära dig:

- Grunderna i att läsa vetenskapliga artiklar.
- Hur man kritiskt värderar originalartiklar.
- Hur man bedömer om resultaten är relevanta.

Metodiken i detta avsnitt är hämtad från Evidensbaserad Medicin, EBM, vilket innebär att man formulerar en frågeställning som besvaras genom systematisk sökning i den vetenskapliga litteraturen, värderar det man funnit och slutligen beslutar om hur man ska handla genom att sammanväga sina resultat med andra faktorer så som patientens önskemål, vårdens organisation, etik, ens egen kompetens etc.

Kursmaterial

- Exempel på vetenskapliga artiklar
- Tips hur man granskar kliniska artiklar
- En checklista för klinisk litteraturgranskning

Arbetsformen

Med utgångspunkt från en klinisk frågeställning ska du träna dig i att kritiskt granska en artikel.

Du ska prova att värdera artikeln utifrån checklistan och artikeln, först på egen hand och sedan gemensamt i gruppen. Checklistan fylls i löpande medan du läser.

Grupparbetet

Gruppindelning enligt tidigare.

Handledarens roll är inte främst att förmedla kunskap till gruppmedlemmarna, utan att få ingång läroprocessen. Eftersom en av målsättningarna med kursen är att deltagarna ska lära sig självstyrt lärande i grupp, utses lämpligen någon av gruppmedlemmarna till att leda gruppen, en ny vid varje gruppssession.

Tips vid kritisk granskning av kliniska artiklar

Vem vänder sig denna sammanställning till?

Idealiskt vore att all sjukvårdspersonal som har det yttersta ansvaret för en speciell behandling (oftast läkare men även sjuksköterskor, sjukgymnaster, arbetsterapeuter etc.) även tränades i den evidensbaserade medicinens fundament: Formulera fokuserade frågor, söka svaren (kostnadseffektivt), granska källorna kritiskt, syntetisera en slutsats och applicera kunskapen på ett omdömesgillt sätt i mötet med den enskilda patienten. Alla skall dock inte bli specialister på kritisk granskning av tillgänglig medicinsk information – gärna evidensbaserad medicin men vi kan inte göra allt grundarbete själva (1 s. 3179). Vi behöver sammanställningar av kritisk granskad vetenskap i form av systematiska översikter eller kommenterade och strukturerade referat av centrala kliniska studier. Om oklarhet råder angående de behandlingar vi själva använder är detta även incitament för egna utvärderingar och studier och detta är faktiskt en del av landstingens ansvar numera (HSU 2000). Denna sammanställning fokuserar på generella problem som kan snedvrída resultat vid kliniska studier och prövningar och vad man bör tänka på vid mötet med läkemedelsindustrin. Det är alltså inte en kurs hur man blir en analytiker vid t.ex. "the Cochrane Collaboration".

Artiklar läser man bl.a. för att vara uppdaterad inom sitt specialområde, enskilda patientfall, uppdrag (t.ex. utarbetande av riktlinjer) eller egen forskning (jmf egna fynd).

Bakgrund och referenser

2000-2002 utgav läkartidningen en serie artiklar om evidensbaserad medicin (EBM) under redaktör Mats Eliasson, öl, med dr, Sunderby sjukhus och tillika projektsamordnare vid Statens beredning för medicinsk utvärdering (SBU) med guidelines hur man kritiskt granskar kliniska artiklar men även kritisk granskning av själva konceptet EBM. Professor Tommy Hansson skriver i förordet följande kloka ord: "Samtidigt som evidensbaserad medicin kan utgöra ett fundament i kunskapsbildningen måste också konstateras att flertalet moment och åtgärder i sjukvården helt saknar underlag för bedömning av vetenskapen" (1). Trisha Greenhalgh, allmänläkare och professor vid University College i London har fokuserat på utbildning av sjukvårdsprofessionen avseende EBM och systematiska översikter men även hur man skall nagelfara kvalitativa studier liksom frågor rörande sjukvårdsadministration och organisation, samt kommunikation och journalistik. Hennes bok "How to read a paper" (2), numera även översatt till svenska, kommer jag att nedan referera till liksom bröderna Curt och Bengt Furbergs bok "Allt är inte guld som glimmar II" – hur man värderar kliniska studier (3). Bengt är docent i klinisk fysiologi, Kungsbacka och Curt Furberg är professor vid Wake Forest University School of Medicine i USA. En annan författare som jag tycker är läsvärd och som fokuserar just på randomiserade kontrollerade kliniska prövningar, s.k. RCT, är den kanadensiske forskaren Alejandro Jadad som i början på 1990-talet arbetade i Oxford men som numera är biträdande chef vid Canadas Cochrane grupp i Hamilton, Ontario. Boken heter på svenska (översatt och fackgranskad av docent i invärtesmedicin Jörgen

Malmquist): Randomiserade kontrollerade kliniska prövningar – En handledning för den som tar del av prövningsresultat (4).

Industrins roll

Läkemedelsindustrins uppgift är att optimera ägarnas intressen, framför allt avseende ekonomi, men även moraliska, etiska och miljömässiga aspekter är viktiga. Att ta fram ett nytt läkemedel kostar i runda tal 10-15 miljarder SEK. Det betyder att läkemedelsstudier **avseende nya läkemedel** i allmänhet är väl genomförda med tillräckligt stora patientmaterial och med medarbetare med hög kompetens avseende medicinsk och statistisk kunskap. Forskningsfusk och snedvridning av data är därför numera ovanligt från industrins sida vad nya läkemedel beträffar, dock inte nödvändigtvis beträffande kombinationer av redan registrerade läkemedel eller nya administrationssätt, inte heller beträffande enskilda forskare. Kolla dock **uppföljningstid** (ofta kort av ekonomiska skäl) samt användandet av **surrogatvariabler** i stället för variabler som patienten upplever som viktiga (se nedan samt ordlistan). Vidare finns inte samma intresse av att ta fram läkemedel för små patientgrupper (sällsynta sjukdomar) som för stora (hypertoni, hjärt-kärl-sjukdomar, osteoporos, demens, etc.) och härvidlag tenderar man ibland att betona obetydliga, men statistisk signifikanta, effekter (se klinisk signifikans och masssignifikans i ordlistan). Det finns också anledning att var skeptisk till megaprövningar (se ordlistan) om man kan misstänka att syftet är att introducera läkemedlet i fråga (till distriktsläkare t.ex.). Vidare finns också en återhållsamhet avseende allvarliga biverkningsrapporter fr.a. avseende storsäljare. När det gäller mötet med läkemedelsrepresentanter är det viktigt att kritiskt kunna ifrågasätta de studier som refereras, inte minst de som handlar om ekvivalensstudier (vårt läkemedel är lika bra som det ni redan använder men billigare eller bekvämare administrationssätt etc.). Även lanseringen av kombinationspreparat bör ifrågasättas om det inte är önskemål från professionen. Det finns också ett relativt intresse att utöka indikationerna för ett eget, relativt billigt, läkemedel om man redan har ett annat, relativt dyrt för den aktuella patientgruppen. Vidare är rapporteringsbias ett problem liksom publikationsbias vid industrifinansierade läkemedelsstudier (se bias nedan samt Box 6.1 sid 88 i ref 2).

Studietyper och vetenskapligt bevisvärde

Olika sammanställningar av redan publicerade studier brukar kallas sekundära studier och övriga primära, där man brukar skilja på experimentella studier och observationsstudier. Man kan även betrakta forskningsområden såsom terapi (farmaka, kirurgiska åtgärder, patientutbildning eller andra interventioner), diagnos (nya diagnostiska tester jmf med någon golden standard), prognos (t.ex. överlevnadsstatistik) eller orsakssamband (är en viss miljöfaktor kopplad till sjukdom? Är en viss organisationsförändring kopplad till förbättring?). Bevisvärdet i den typ av studie man granskar varierar. Härvidlag brukar systematiska översikter, guidelines (Cochrane, SBU etc.) och metaanalyser komma först, följt av randomiserade kontrollerade prövningar (RCT) i sin tur följt av observationsstudier.

Systematisk review och metaanalys

Randomiserad prospektiv studie - med signifikant resultat

Randomiserad prospektiv studie - med misstänkt men inte signifikant resultat

Kohortstudie

Fall-kontroll-studie

(Tvärsnittsstudie)

Förändring pga. intervention vid statistisk processkontroll (SPC)

Fallbeskrivning

Förutom ovanstående studietyper, som är kvantitativa, finns kvalitativa studier (se ordlistan).

Vad **evidensstyrka för slutsatser** beträffar indelas vetenskapligt underlag i fyra nivåer (1):

A) Starkt: Minst två studier med högt bevisvärde eller god systematisk översikt.

B) Måttligt: En studie med högt och minst två med medelhögt bevisvärde.

C) Begränsat: Minst två studier med medelhögt bevisvärde.

D) Otillräckligt: Annat underlag: Vad?

Övergripande frågor vid artikelgranskning

- Tillför studien något nytt (större, längre, bättre, annan population?).
- Tillför studien nya data för t.ex. metaanalys?
- Bemöter studien tidigare kritik eller finns fortfarande tveksamheter (t.ex. bland allmänhet eller politiker).
- Vilka resultat visade studien och är dessa tillförlitliga?
- Var uppföljningstiden tillräckligt lång? Ofta krävs 5-10-15 års uppföljning som ingen orkar vänta ut. Det är dock viktigt att fundera över detta, t.ex. vid behandling av benskörhet eller andra förebyggande åtgärder.
- Var studiepopulationen tillräckligt stor. Läkemedelsprövningar brukar alltid ha tillräckligt stor studiepopulationen (se ordlistan; klinisk signifikans och masssignifikans), möjligen undantaget s.k. ekvivalensstudier, dvs. studier där man vill visa likhet med någon redan etablerad (men kanske dyr eller krånglig) behandling. Genom att ha en förhållandevis liten sampelstorlek kan man visa att inga signifikanta skillnader föreligger. Här är det den medicinska professionen (och inte läkemedelsindustrin eller någon statistiker) som avgör vad som är medicinsk acceptabelt.
- Hur representativa är patienterna i kliniska prövningar? Mer (eller mindre) sjuka? Annan socialklass, kön eller ålder? Comorbiditet? Exklusionskriterier av typ rökning, alkohol etc.)? **Prövningspatienter kan vara en exklusiv grupp.**

- Hur viktig är frågan för mina patienter? Det är bra att veta att patientens preferenser nästan alltid är följande: att *må bättre*, att *minska risken för framtida icke-dödliga komplikationer*, att *förlänga livet*. Annat får anses vara surrogatvariabler.

Patientaspekten är intimt kopplad till effektmåtten och de etiska frågorna. Jag tar därför upp surrogatvariabler och forskningsetiska frågor här innan jag går vidare med bias.

Surrogatvariabler: För att bedöma patientrelaterade variabler används ofta variabler som i sin tur är kopplade till framtida risker eller prognos. Sådana kan vara bentäthet i st.f. frakturincidens, ventrikulära extraslag på långtids-EKG i st.f. arytmidöd, arteriell hypertension i st.f. stroke, hyperkolesterolemi i st.f. infarktincidens etc. Sådana variabler kallas surrogatvariabler (se ordlistan). T.ex. ledde fluoridbehandling av benskörhet till ökad bentäthet men dessvärre också till ökad frakturincidens (5) och de ventrikulära extraslagen minskade på långtids-EKG men dödligheten ökade med flecainid (6). *Det är viktigt att skilja äkta vara från surrogat.*

Forskningsetiska aspekter: Forskningsetiska frågor är numera oftast inte något stort problem vid artikelgranskning i västvärlden (t.ex. EU, USA, Canada, Australien, Nya Zeeland) beroende på den strikta forskningsetiska bedömningen som föregår studierna här. Studier från diktaturer måste alltid bedömas med stor försiktighet (se även publikationsland-bias).

Läs abstract först

Börja med att läsa abstract i sin helhet för att om möjligt snabbt avgöra om artikeln överhuvud taget är värd att läsa. Är det rätt studiedesign, patientgrupper eller metod för att besvara frågan. I abstract finns oftast uppgifter om hur behandlingen har givits, hur kontrollgruppen har behandlats, hur effekten har mätts (det finns t.ex. anledning att ifrågasätta studier där doktorn bedömt patientens klåda, yrsel eller sömnstörning), vilken sjukvårdsmiljö som studien har genomförts i. Redan här kan man alltså avgöra om artikeln är värd att läsa. Om inte, lämnar man den och har därmed sparat avsevärd tid. Om Du nu vill läsa vidare kan vidstående **granskningsmall** kanske vara behjälplig (Sammanfattande frågor vid artikelgranskning).

Bias

Kraven på det kritiska läsandet kommer att variera beroende på vilket forskningsområde eller vilken studietyp man granskar. I de följande kommer jag att lägga fokus på de snedvridningar (bias) av sanningen som man kan råka ut för fr.a. vid studier med potentiellt högt bevisvärde av typ randomiserade kontrollerade prövningar (RCT). Mycket av detta är dock generaliserbart till medicinsk forskning i allmänhet.

Randomiserings - och selektionsbias: Randomisering bör göras centralt eller med ogenomskinliga kuvert (concealed allocation). Om randomiseringen är perfekt men inte blindningen, kan prövaren påverka deltagandet antingen genom att

tillämpa exklusionskriterierna selektivt eller genom att beskriva prövningen för patienten så att vissa patientgrupper väljer att inte delta (se sid 53 ref 4). Beskrivningen av randomiseringsförfarandet är viktigt och blindningen ofta begränsad.

Utförandebias: Angående blindning, se ordlistan. Dubbelblind betyder oftast trippelblind, behandlande, patient, utvärderare. Ofta svårt men bör eftersträvas, fr.a. vid subjektiva effektmått, men inte lika viktigt som randomiseringsförfarandet. Även kirurgisk intervention kan (och bör om möjligt) blindas. Man kan lätt förstå att deltagare som får read på att de fått placebo kan känna sig besvikna och ovilliga att framföra att symptomen minskat eller försvunnit. På likartat sätt kan för prövarna, medvetet eller omedvetet, bli mer positiva avseende bedömningar och journalanteckningar om de vet att patienten fått aktivt läkemedel liksom omhändertagande i övrigt.

Bortfall: Bortfall kan ske av uppenbara eller mindre uppenbara skäl, t.ex. biverkningar av läkemedlet, bristande motivation hos patienten eller dålig compliance av andra skäl, kliniska skäl (annan sjuk, graviditet etc.), bortfall vid uppföljning (pga. flyttning, gruppbyte (se intention to treat och analys per protokoll i ordlistan) etc.), dödsfall, eller det faktum att patienten bara försvinner. I en två års uppföljningsstudie avseende sekundärprofylax mot plötslig död efter hjärtinfarkt med Anturane (sulfinpyrazone, ett trombocytaktivt medel) bedömdes 71 av 1629 patienter **felinkluderade i efterhand** eftersom patienterna inte ansågs uppfylla kriterierna för att ingå i studien. Oturligt nog för prövarna observerade editorn i New England Journal of Medicine att $10/38=26\%$ av de exkluderade hade avlidit i behandlingsgruppen medan motsvarande siffra i placebogruppen endast var $4/33=12\%$. Denna subjektiva utslutning påverkade hela studieresultatet. Slutsatsen av detta är att man bör vara återhållsam med att i efterhand utsluta patienter som en gång inkluderats. Om misstag begåtts vid inkluderingen är det sannolikt bättre att leva med detta misstag än att exkludera patienterna efter det att behandlingsresultatet blivit känt. **Det är ett dåligt tecken när patienter bara försvinner.**

Ett exempel på felaktig bortfallsanalys är följande: I the coronary heart project (8) jämfördes flera lipidsänkare bl.a. klofibrat med placebo avseende den totala dödligheten under en 5-årsperiod efter hjärtinfarkt. För att se om klofibrat var bättre om man verkligen tog preparatet (good compliers) jmf med om man inte gjorde det (bad compliers) undersökte man dessa subgrupper. Man kunde då konstatera att dålig compliance var kopplat till sämre överlevnad (25% av bad compliers dog jmf med 15% av good compliers). Detta skulle kunna tolkas till klofibrats fördel, enär de som tar preparatet har bättre överlevnadschans. Det visade sig dock vid senare analys av placebogruppen att motsvarande siffror var 28% för bad compliers och 15% för good. Uppenbarligen spelar det ingen roll om man tar klofibrat eller sockerpiller. Orsaken är att bad compliers generellt sett har högre dödlighet eftersom de tillhör en selekterad grupp (t.ex. missbrukar alkohol och tobak, annan socialgruppsstillhörighet etc.).

Bedömnings-och analysbias: Minskas genom olika grad av blindning (se ordlistan; ang. bedömningsbias under studietiden se utförandebias). Dessutom att man strikt håller sig till ursprungsprotokollet avseende grupper och bedömningsvariabler. Det är i detta skede som icke-planerade subgrupperingsanalyser tas fram liksom nya effektvariabler och man kan optimera

tiden så att uppföljningsresultatet förefaller bli så bra som möjligt. Se även rapporteringsbias.

Presentationbias: Vid presentation av data kan författarna ibland lockas att överdriva skillnader genom att hugga av skalor i diagram (trunkering), att redovisa relativa riskminskningar i st.f. absoluta och att undvika att redovisa antalet personer som måste behandlas för att *en* skall undgå någon specificerad händelse (infarkt, död, fraktur, recidiv etc.; se number needed to treat, NNT, i ordlistan). Inte sällan ses även diagram med förändring ”skillnad efter-före behandling” på y-axeln plottat mot ”före behandling” på x-axeln, vilket alltid leder till statistiskt signifikanta samband men som är en slumpmässig, matematisk oundviklighet (s.k. pseudokorrelation). Slutsatsen kan då felaktigt bli: Ju större x desto större effekt av interventionen.

Rapporteringsbias: Har den vetenskapliga frågeställningen specificerats i förväg? Var det verkligen den frågan man senare besvarade? Somliga ställer inte frågan förrän de vet svaret. Rapporteringsbias anses vara ett större problem än publikationsbias. I studier som faktiskt publiceras rapporteras företrädesvis de mest positiva resultaten (se publikationsbias nedan). För att genomskåda selektiv rapportering måste protokollet vara känt för läsaren vilket oftast inte är fallet. Läkemedelsverket har dock tillgång till sådana och vid jämförande studier (skillnad mellan vad som avsågs att rapportera och vad som de facto rapportades) kan man ofta se mycket överskattade resultat (9).

Publikationsbias: Negativa studier är mindre intressant att fullfölja. Negativa studier är svårare att få publicerade. Publikationsbias är ett stort problem vid metaanalys och man försöker nu minska denna genom att göra det obligatoriskt att anmäla alla påbörjade medicinska studier till någon central enhet. Se också upp för halvtidsresultat. Celecoxib longterm safety study (CLASS) ville visa att Cox-2-hämmaren Celebra hade mindre biverkningar i form av gastrointestinala blödningar jämfört med icke-selektiva Cox2-hämmare (ibuprofen eller diklofenak (Voltaren)). Halvtidsresultatet (6 mån) visade mindre ulcuskomplikationer men 12 månadersresultatet visade ingen skillnad. Man valde då att publicera 6 månadersuppföljningen (inte enligt studieupplägget) och inte 12 månaders (10). En form av publikationsbias är s.k. *fördröjningsbias*, vilket innebär att studier med negativa resultat publiceras senare (ca dubbelt så lång tid att bli publicerade jämf med studier med positivt resultat). I samma genre kan även bias vid ”potentiellt genombrott” räknas, där journalister och Internetredaktörer systematiskt väljer ut, överskattar och ger spridning åt provningsresultat som går i en riktning som de anser vara intressant och ”glömmer bort” att flera tidigare studier visat motsatsen.

Publikationsland-bias: Det har visats att forskare i vissa länder enbart publicerar positiva resultat, som t.ex. RCT som undersöker akupunktur i Kina, Japan, Hong Kong och Taiwan (11).

Referenser

1. Evidensbaserad medicin EBM, särtryck av en serie artiklar ur läkartidningen 2000-2002. Ed. Mats Eliasson.
2. Greenhalgh, Trisha. How to read a paper – the basics of evidenced based medicine, sixth impression, BMJ Publishing group, London, ISBN 0-7279-1139-2, 2000.
3. Bengt Furberg, Curt Furberg. Allt är inte guld som glimmar – II – Hur man värderar kliniska studier. BTJ Tryck AB, Lund, 2001. ISBN 91-631-1241-8.
4. Alejandro Jadad. Randomiserade kontrollerade kliniska prövningar – En handledning för den som tar del av prövningsresultat. Studentlitteratur, Lund, 2000. ISBN 91-44-01578-X.
5. Riggs BL et.al. Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis. N Engl J Med 1990; 322:802-9.
6. Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH, et.al. Mortality and morbidity in patients receiving encainide, flecainide or placebo. N Engl J Med 1991; 324:781-8.
7. Anturane Reinfarction Trial Research Group. Sulfipyrazone in the prevention of sudden death after myocardial infarction. N Engl J Med 1980; 302: 250-6.
8. Coronary Drug Project Research group. Influence of adherence to treatment and response of cholesterol on mortality in the Coronary Drug Project. N Engl J Med 1980; 303; 1038-41.
9. Melander H. selektiv rapportering - ett större problem än selektiv publicering? Läkartidningen nr 4, 2005 volym 102: 22-225.
10. Ledare i BMJ 2002;324:1287-8. Läkartidningen 40: 2002, vol99, s 3934.
11. Vickers A, Goyal N, Harland R and Rees R. Do certain countries produce only positive results? A systematic review of controlled trials. Controlled Clin Trials 1998; 19:159-66.

Sammanfattande frågor vid artikelgranskning (modifierat från FORSS´ FoU-kurs och ref 1).

Författare: Är det tydligt vilka som utfört studien. Finns ekonomiska eller andra intressen? Kopplingar till industrin? Vilket land kommer artikeln ifrån? (se publikationslands-bias).

Frågeställning: Är frågeställningen (syftet) väl beskriven och väl definierad? (Se rapporteringsbias).

Introduktion: Har frågeställningen satts i relation till aktuell kunskap? Är referenserna aktuella? (se även genombrottsbias).

Studiedesign: Vilken studiedesign (typ av studie) har valts? Är studiedesignen väl vald med avseende på frågeställningen?

Material: Är studiepopulationen väl beskriven? Urvalskriterier? Är grupperna jämförbara? Är studiepopulationen unik (titta på exklusionskriterierna)? Hur har randomiseringen skett (om sådan gjorts)? (se selektions-och urvalsbias). Är bortfallet noga beskrivet? Skevhet? Storlek? (se bortfallsbias)

Metod: Har man använt en validerad metod? Är metoden relevant, väl beskriven och reproducerbar? Är studien blindad? Vem är den i så fall blindad för? Om inte blindad, hur har man minskat risken för behandlingsbias (studiegruppen mer omtanke/bättre omhändertagande etc.)?

Resultat: Vilket är studiens huvudresultat? Svarar resultatet på den primära frågeställningen (syftet)? (se rapporteringsbias). Har resultaten bedömts blint? Kan man misstänka att resultatmått ändrats? Har subgrupperingar tillkommit i efterhand? Är uppföljningstiden planlig. Har exkludering av patienter även gjorts i efterhand? (se bedömnings- och rapporteringsbias). Hur har man redovisat eventuella skillnader mellan grupper? Har "intention to treat"- analys gjorts, dvs. analyserades patienterna i den grupp de ursprungligen randomiserades till (se ordlistan)? Hur stor var behandlingseffekten (vid interventionsstudie). Är ev. diagramskalor trunkerade (huggna) för att överdriva skillnader? Redovisas kliniskt relevanta effektmått, t.ex. "number needed to treat" (NNT)? (se ordlistan samt redovisningsbias).

Diskussion: Hur har man bedömt resultaten i relation till tidigare kunskap? Hur har man diskuterat ev. gruppskillnader? Är eventuella statistiskt signifikanta skillnader även kliniskt signifikanta? Hur kan eventuella bortfall ha påverkat resultatet? Har detta diskuterats? Vid observationsstudier tillkommer även frågor ang. vilka confounders som man tagit hänsyn till och tänkbara källor till felklassificering vid fall-kontrollstudier. Är resultaten generaliserbara (gäller resultatet även mina patienter)? Diskuteras ev. begränsningar? Är slutsatserna väl avvägda?

Bedömning: Vad anser Du om studiens resultat? Håller Du med författarna? Trovärdiga? Tillförlitliga? Begränsningar i tillämplighet? Är studien relevant för Dig? Anser Du att resultaten av denna studie borde leda till några åtgärder? I så fall vilka?

Liten ordlista

EBM (Eng. evidenced based medicine), Evidensbaserad medicin: Varje åtgärd i sjukvården skall vara baserad på bästa möjliga tillgängliga kunskap.

Studietyper och prövningar och därtill hörande begrepp

Kvantitativ studie	Studie som bygger på analys av siffror, inklusive t.ex. kategoridata och proportioner.
Kvalitativ studie	Studie som bygger på analys av text, t.ex. intervjuer som skrivs ned.
RCT	(Eng. randomized controlled trial), randomiserad kontrollerad klinisk prövning: Undersökning där personer på ett slumpmässigt sätt fördelas till en eller flera kliniska åtgärder. Kontrollerad betyder att man försöker hålla grupperna så lika som möjligt avseende allt utom de kliniska åtgärderna som avses studeras.
Observationsstudie	Icke-experimentella studier, ofta benämnda epidemiologiska studier.
Explanatorisk prövning	(Eng. explanatory trial): Prövning som svarar på frågan: Har interventionen någon effekt (efficacy) oftast jmf med placebo. Typiskt är att utvalda grupper inkluderas (patienter inom snävt åldersintervall som bara har den aktuella sjukdomen och som kan förväntas följa ordinationer etc.). Det snarlika begreppet efficacy-prövning (eng. efficacy-trial) finns också.
Pragmatisk prövning	(Eng. management trials): Avser att belysa och beskriva alla konsekvenser av interventionen, goda och dåliga, under omständigheter som överensstämmer med vardagssjukvårdens. Här är urvalet mindre strikt jmf explanatoriska prövningar för att få med "vanliga" patienter. Här får man svar på frågan hur effektivt läkemedlet är i vardagssjukvården (effectiveness). Det snarlika begreppet effectiveness-prövning (effectiveness trial) finns också.
Fas I-prövning	Första experimenten på människa. Ett fåtal forskningspersoner.
Fas II-prövning	Prövningar på små grupper av patienter med aktuell sjukdom, med hänblick på efficacy avseende dosstorlek och administreringsintervall etc.
Fas III-prövning	Oftast effectiveness-prövningar med större patientgrupper och kontrollgrupper. Oftast RCT.

Parallellgruppsprövning	Varje grupp blir bara föremål för en av behandlingarna. Möjliggör jämförelser mellan grupperna.
Överkorsningsprövning	(Eng. crossover trial): Varje grupp blir föremål för båda behandlingarna (kommer nästan bara i fråga vid studier med två interventionsalternativ, t.ex. aktivt läkemedel och placebo). Möjliggör jämförelser både mellan och inom grupperna.
Faktoriell prövning	Studie vars avsikt är att bedöma kombinationer av interventioner, t.ex. jämförelse mellan de fyra grupperna som behandlas med analgetika 1, analgetika 2, analgetika 1+2 samt placebo. Man kan då mäta effekterna av de enskilda preparaten samt kombinationen av preparaten och ställa dessa emot varandra och mot placebo.
"Megaprövning"	Multicenterprövning med 1000-tals forskningspersoner i syfte att uppnå hög generaliserbarhet och hög statistisk styrka. Ibland även enbart av marknadsekonomiska skäl (dvs. reklam; som förhoppningsvis i så fall stoppas i etikprövningsnämnden (EPN))).
Bias	Snedvridning: Felaktigheter i vetenskapliga resultat pga. något systematiskt fel beträffande t.ex. randomisering, insamling av data, bortfall, bearbetning, analys, redovisning, rapportering och publikation av studieresultat.
Blindning	Följande varianter på blindning finns: 1) Ingen blindning (öppen studie: Kan vara en kontrollerad randomiserad prövning (RCT) men där alla som är inblandade i prövningen vet vilken grupp som behandlats med vad; de flesta kirurgiska studier eller kirurgi mot läkemedel, koststudier etc. tillhör denna grupp). 2) Enkelblindning (endera forskningspersonerna eller prövarna). 3) Dubbelblindning (både forskningspersonerna och prövarna), 4) Trippelblindning (som dubbel plus t.ex. de som bedömer och registrerar eller de som tar hand om databearbetning och statistisk analys). 5) Kvadrupelblindning: (som trippel men även författarna. Dessa skriver på gruppnivå och får först efter rapportfärdigställandet reda på grupptillhörigheterna).
Effektått	(Eng. outcome measures): På förhand bestämda variabler (effektvariabler) och tillhörande mätmetoder. T.ex. antalet händelser av ett visst slag under studiens gång (infarkt, fraktur, död etc.), (eng. event, endpoint), kontinuerliga variabler (blodtryck,

	överlevnadstid etc.) eller andra variabler (enkätsvar etc.).
Surrogatvariabler	Variabler med väl dokumenterad koppling till för patienten betydelsefulla tillstånd eller risker men som inte patienten själv märker. Hypertension är kopplad till strokerisk, minskad bentäthet är kopplad till frakturrisik etc. Effekt på surrogatmått har i sig inte något värde för patienten och det är viktigt att alla behandlingar förr eller senare utvärderas avseende verkliga end-points som risk för död, infarkt, fraktur, ökad livskvalitet, ökad livslängd, smärtlindring etc. Blodtryck är ett av få surrogatmått som anses oberoende kopplat till risk.
Forskningspersoner	(Eng. subjects), deltagare i forskningsprojekt (benämningen försökspersoner skall inte längre användas).
Intervention	Kliniska åtgärder: Kan utgöras av förebyggande åtgärder, diagnostiska tester, screening-program, behandlingsåtgärder.
Compliance	Följsamhet, dvs. i hur grad forskningspersonen följt anvisningarna.
Monitorering	Kontrollerande åtgärder som vidtas för att säkra att prövningen utförs strikt enligt planen, att testresultatet bokförs korrekt etc. (icke att förväxlas med kontrollerad i RCT, se d.o.)

Statistisk och analys

Intention to treat	Typ av utfallsanalys som innebär att alla forskningspersoner inkluderas i analyserna som del av de grupper till vilka de blivit randomiserade, oavsett om de fullbordat prövningen eller inte eller om de bytt grupp. En person som t.ex. randomiserats till medicinsk behandling men som under studiens gång opereras (dvs. övergår till den kirurgiska gruppen) kvarstår i analysen som medicinskt behandlad. Orsaken är att man hellre undviker randomiseringsbias på bekostnad av lägre styrka (power). Ofta redovisas dock även det "verkliga" resultatet, s.k. analys per protokoll (efficacy-study).
Analys per protokoll	Efficacy-study. Se Intention to treat.
H ₀ :	Nollhypotesen. Vid hypotesprövning betyder H ₀ att allting är som det alltid har varit. T.ex. när man behandlar en grupp patienter med ett nytt läkemedel så sker ingen förbättring.
Typ-I fel (α)	Om H ₀ är sann men resultatet felaktigt talar för det motsatta, t.ex. att studieläkemedlet har effekt har man gjort ett typ-I fel. Detta fel försöker man hålla

	<p>lågt, ett vanligt mått är <5 %. Sådana fel kan vara dyrt för samhället och orsaka helt onödiga utgifter, förutom ev. patientlidande.</p>
Typ-II fel (β)	<p>Om H_0 är falsk och studieresultatet inte kan påvisa någon skillnad har man gjort ett typ-II fel. $(1-\beta)$ brukar kallas för studiens power (styrka). Man brukar nöja sig med en power på 80-90 %, dvs. ett β-fel som är dubbelt eller fyrdubbelt större än α. Detta fel har inte så stora samhälleliga ekonomiska konsekvenser, åtminstone inte på kort sikt (men kan orsaka patientlidande; jmf asbest, hormoslyr, rökning som det tog decennier att koppla till sjukdom trots befintliga resultat). Skulle ett bra läkemedel undgå upptäckt av behandlingseffekt och företaget ändå tror på preparatet får väl studien göras om på ett större material.</p>
p-värde	<p>Ett tal mellan 0 och 1 som anger sannolikheten att göra ett typ-I-fel. P-värdet säger inget om typ-II felet, vilket är av speciellt intresse när H_0 accepteras. Typ-II felet (power) är alltså viktigt fr.a. vid påståenden av typ: Här finns alltså ingen statistiskt signifikant skillnad, hurra!</p>
Oönskade effekter	<p>Adverse events: T.ex. oönskade biverkningar av studieläkemedlet.</p>
Klinisk signifikans	<p>Effektskillnad som kan anses vara kliniskt betydelsefull. Denna är naturligtvis beroende på sjukdomens svårighetsgrad och en fråga för den medicinska professionen, inte för statistiker eller läkemedelsföretag. Om man måste behandla 20 snuvor för att en patient skall få kortare sjukdomsduration är det inte samma sak som att behandla 20 personer för att rädda ett liv.</p>
Prospektiv	<p>Framåtblickande. Alla RCT är prospektiva studier.</p>
Retrospektiv	<p>Tillbakablickande. En registerstudie är exempel på en retrospektiv studie.</p>
Sampelstorlek	<p>(Eng. sample size): Forskningspersonernas antal (n).</p>
Prövare	<p>(Eng. investigators), undersökare, kliniska forskare: De personer som utformar undersökningen, utför interventioner, bedömer och analyserar resultaten.</p>
Masssignifikans	<p>Statistisk men kliniskt tveksam eller betydelslös signifikans beroende på stora patientmaterial.</p>
Randomisering	<p>Slumpmässig fördelning. I strikt bemärkelse endast om alternativen är fullständigt dolda (för att undvika randomiseringsbias). Sålunda är ett allokeringsförfarande med varannan patient till</p>

	studie- resp. kontrollgrupp inte en strikt randomisering vid en prospektiv studie.
Sensitivitet	Ett mått mellan 0 och 1 på hur känsligt ett test är att plocka ut personer med en viss sjukdom ur en population, dvs. kvoten mellan sjuka med positivt test och alla sjuka. Sensitiviteten är således <u>inte</u> beroende av prevalensen av sjukdomen i undersökningspopulationen eftersom endast de sjuka betraktas. Tester där man är intresserad av att utesluta en viss sjukdom skall alltså ha hög sensitivitet.
Specificitet	Ett mått mellan 0 och 1 på hur specifikt ett test är, dvs här gäller det att hålla nere antalet friska med positivt utfall, alltså de falskt positiva. Man får då tänka omvänt, dvs. hur bra är testet på att välja ut friska ur den friska delen av populationen, dvs. kvoten mellan friska med negativt test och alla friska. Specificiteten är, liksom sensitiviteten, således inte beroende av prevalensen av sjukdomen i undersökningspopulationen eftersom endast de friska betraktas.
NPV	(Eng. negative predicted value), negativt predicerat värde. Om personens test är negativt, vad är då sannolikheten för att han är frisk, dvs. kvoten mellan friska med negativt test och alla med negativt test. Här tittar man alltså endast på de som har negativt testresultat men som består av både sjuka och friska (om inte testet är 100-procentigt säkert för då finns ju inga sjuka med här). NPV är alltså beroende på prevalensen av sjukdomen och blir då olika i olika populationer. På "packsedeln" till ett sådant test kan alltså sensitivitet och specificitet stå angivna men inte NPV och PPV.
PPV	(Eng. positive predicted value), positivt predicerat värde. Som NNV men endast avseende positivt test.
NNT	(Eng. number needed to treat). Inversen av den absoluta riskminskningen (eng. absolute risk reduction; ARR). Obs att industrin hellre talar om relativ riskminskning (RRR) som i allmänhet är betydligt större än ARR. T.ex. om man har en riskreduktion från 8% i kontrollgruppen till 6% i behandlingsgruppen har man en RRR på $(0.08 - 0.06)/0.08 = 25\%$ medan $ARR = 2\%$ och $NNT = 1/0.02 = 50$, dvs. 50 patienter måste behandlas för att 1 person skall slippa den aktuella händelsen (t.ex. infarkt eller vad det nu kan vara).
NNH	(Eng. number needed to harm). Inversen av den absoluta riskökningen (eng. absolute risk increase;

ARR). Här finns varierande definitioner beroende på typ av risk. En definition är helt enkelt NNH= NNT om detta är negativt, dvs. då är effektvariabeln densamma för de två talen (t.ex. död, hjärtinfarkt, stroke etc.). Riskökning vid NNH kan även beskrivas som olika typer av biverkningar, oftast irreversibla – i så fall är det viktigt att man redovisar dessa om NNH används. Om exempelvis trombolys vid hjärtinfarkt minskar risken för död från 12 till 9 %, är $NNT=1/0.03=33$ patienter. Samtidigt ökar risken för hjärnblödning med 1 %, dvs. $NNH=1/0.01=100$. Det är naturligtvis viktigt att väga dessa tal mot varandra liksom följderna av de olika riskerna.

Oddsquot (OR)

(Eng. Odds ratio): Odds definieras som $p/q = p/(1-p)$ där p är sannolikhet för t.ex. sjukdom och q är sannolikhet för inte sjukdom ($1-p$; obs detta p har inget att göra med p -värde vid signifikansbedömning). Oddsquoten (OR) = odds i studiegruppen/odds i kontrollgruppen. Ett OR på 0.5 betyder då hälften så stor risk för sjukdom i studiegruppen jmf kontrollerna. Man brukar ange OR som väntevärde och 95 %-igt konfidensintervall, t.ex. 0.50 (0.32-0.85). Om detta intervall överlappar värdet 1 är skillnaden inte statistiskt signifikant på 0.05-nivån. Riskquot (RR) och OR är approximativt lika för sällsynta händelser $p<5-10\%$. Obs, i fall-kontrollstudier kan OR men inte RR beräknas, eftersom prevalensen är okänd.

Relativ risk (RR)

(Eng. relative risk), detsamma som riskquot (eng. risk ratio), RR. Risk (R) definieras som sannolikheten för att någon händelse sker under viss tid, dvs. $R=p$ (samma p som i OR ovan). RR kan då vara kvoten (R i studiegruppen)/(R i kontrollgruppen), dvs. $R(s)/R(k)$. Relativ riskminskning (RRR)=[$R(k)-R(s)$]/ $R(k)$ och absolut riskminskning (ARR) = $R(s)-R(k)$. I epidemiologiska termer får man ersätta minskning med skillnad, men då talar man inte om RRR utan bara RR (och OR).

Rate

Observera att rate är en hastighet, dvs. förändring/tidsenhet. Ratequot (eng. rate ratio) är således en kvot mellan hastigheter (t.ex. kvoten mellan hur fort en patientgrupp med cancer dör med behandling och motsvarande grupp utan behandling).

Logistisk regression

Icke-linjär, ofta multipel, regression där man använder funktionen $\text{logit}(p)=\ln(p/q)$, dvs. $\log(\text{odds})$, $q=1-p$. Det finurliga är att

regressionskoefficienterna är logoddskvoter, dvs. om man beräknar exponentialfunktionen av dessa får man OR. Den beroende variabeln är dikotom (typ sjuk/frisk).

Coxregression

Icke-linjär, ofta multipel, regression där man beräknar s.k. hazard rate ratios (HR), dvs. kvoter liknande logoddskvoter men mellan lutningarna på överlevnadskurvor. Den beroende variabeln är dikotom (typ levande/död).
Regressionskoefficienterna är $\log(\text{HR})$.