
Vad säger oss ”efterstudier” och metaanalys om effektiviteten av interventioner för mental hälsa?

JOHN P.A. IOANNIDIS

Metaanalys har accepterats som den högsta nivån i evidenshierarkin och den kan erbjuda kvantitativa metoder också för att bedöma systematiska fel i klinisk forskning. Här diskuterar vi vissa empiriska tillämpningar inom området mental hälsa. För det första lägger vi fram evidens för en suboptimal balans vid rapporteringen av för- och nackdelar i kliniska prövningar vid psykisk ohälsa. För det andra framläggs evidens om i vilken mån små prövningar ger samma resultat som stora. Stora kliniska prövningar är mycket ovanliga vid psykisk ohälsa, och kan, då de genomförs, ibland ge andra resultat än små prövningar. Till sist framläggs empiriska data om huruvida evidensen förändras med tiden. Tillämpningen av rekursiv kumulativ metaanalys tyder på att tidiga prövningar kan ge fördelaktigare resultat än senare prövningar, speciellt vid läkemedelsterapi.

VAD KAN METAANALYS ERBJUDA EVIDENS-BASERAD MEDICIN?

Evidensbaserad medicin lägger stor vikt vid att kritiskt granska samlad information inom hälso- och sjukvården samt att åstadkomma en systematisk syntes av data från olika studier av samma kliniska spörsmål eller forskningsfråga. Inom flera system för rangordning av evidenshierarkin har metaanalys under det senaste decenniet nått en topposition av utomordentlig betydelse.

Metaanalys omfattar en samling kvantitativa metoder som syftar till att på ett gediget matematiskt sätt integrera tillgänglig information om en specifik fråga. Ändamålet med metaanalys ses ofta som ett försök få sammanfattade uppskattningar av olikartad information, men det är antagligen en missuppfattning om vad denna forskningsmetod omfattar. Metaanalys kan visserligen förbättra styrkan i att upptäcka skillnader eller att övertygande visa att det inte finns skillnader mellan de jämförda modaliteterna. Att upptäcka, bedöma och förklara olikhet (heterogenitet) är dock antagligen ett intressantare mål för metaanalys än den enkla datasyntesaspekten

(1). På senare tid har det utvecklats flera kvantitativa metoder som gör det möjligt att analysera källorna till heterogenitet och systematiskt fel. Detta tillvägagångssätt kan öka vår insikt om kliniska interventioner, klarlägga problem med tidigare forskning och antyda prioriteringar med avseende på framtida forskning.

I de medicinska vetenskaperna har metaanalys tillämpats främst för datasyntes från randomiserade kliniska prövningar. Men det finns även ett ökande intresse för till-

FÖRFATTAREN

John P.A. Ioannidis är professor och chef för institutionen för hygien och epidemiologi vid medicinska fakulteten vid universitetet i Ioannina, Grekland och biträdande professor i medicin vid Tufts University School of Medicine, Boston, USA. Han är intresserad av evidensbaserad medicin samt klinisk och molekylär epidemiologi.

lämpningar som går ut på att sammanställa och undersöka heterogenitet från icke-randomiserade epidemiologiska data, molekylärmedicinsk information, diagnostiska tester eller prognosfaktorer. Dessa nya tillämpningar innebär egna utmaningar då det gäller att genomföra metaanalys och tolka resultaten.

Termen "efterstudier" är antagligen också missvisande för metaanalys och liknande forskningsmetoder. De data som analyseras är naturligtvis oftast retrospektiva, men denna regel har flera undantag. Prospektiva metaanalyser vinner ökat stöd som ett lockande tillvägagångssätt för att få fram evidens i stor skala. Också retrospektiva metaanalyser är för övrigt "primärforskning" så till vida, att de kan och bör ställa upp sina egna, klart specificerade forskningsobjekt och mål. En stor fördel med en metaanalys är att tillvägagångssätt och protokoll kan vara ytterst standardiserade, i nästan lika hög grad som vid en randomiserad prövning. Dessutom kan systematiska översikter och metaanalyser visa på viktiga brisfälligheter inom den medicinska litteraturen som kräver korrigering. Här följer några exempel på sådana tillämpningar inom området mental hälsa.

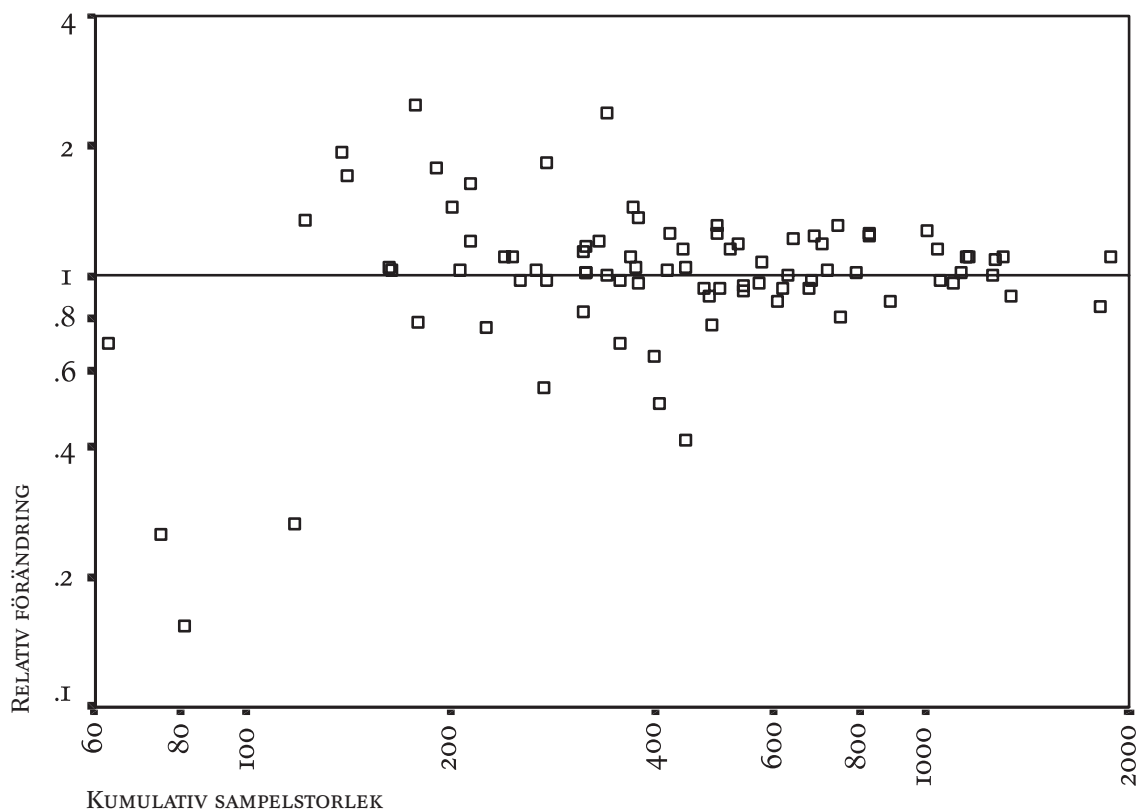
BRISTFÄLLIGHETER OCH SYSTEMATISKA FEL I RAPPORTERAD KLINISK FORSKNING: FÖR- OCH NACKDELAR

Granskningen av kliniska studier inom psykisk ohälsa har visat på flera frågor i planeringen, rapporteringen och tolkningen som kräver uppmärksamhet. Första rangens problem innefattar små sampelstorlekar, fragmentering av litteraturen och användning av hundratals skalor för att mäta utfallet, av vilka många aldrig har validerats. Det ser också ut som om rapporteringsstandarder negligeras. Förhoppningsvis kommer en omfattande tillämpning av Consolidated Standards for Reporting of Trials (CONSORT) att förbättra denna situation i framtiden (2). Vissa faktorer i samband med rapporteringen uppvisar stora brister och kräver speciell uppmärksamhet. Speciellt verkar det finnas en obalans i hur effekt och skadeverkningar behandlas i den medicinska litteraturen. Vi bedömde nyligen 142 randomiserade prövningar från registret PsiTri (tillgängligt på <http://psitri.stakes.fi>) över kontrollerade prövningar inom psykisk ohälsa (3). Vi analyserade dessa prövningsrapporter med avseende på tillräcklighet av och den relativa tonvikten på säkerhetsfrågor i innehållet. Av de 142 prövningarna hade 25

minst 100 randomiserade försökspersoner och minst 50 försökspersoner i var studiearm. I fråga om läkemedelsprövningar hade bara 21,4 procent adekvat rapportering av kliniska skadeverkningar och bara 16,5 procent hade adekvat rapportering av laboratoriebestämmd toxicitet, medan 32,0 procent rapporterade både antalet och orsakerna till bortfall på grund av toxicitet i vardera armen. Vid läkemedelsprövningar offrades i medeltal en tiondels sida i resultatrapporterna på säkerhet och 58,3 procent ägnade mera utrymme åt författarnas namn och positioner än åt säkerhet. Ingen av prövningarna av icke-läkemedelsinterventioner (n = 39) hade tillräcklig eller ens delvis tillräcklig rapportering av vare sig kliniska skadeverkningar eller av laboratoriebestämmd toxicitet. Det fanns ingen undergrupp som gav betydande utrymme åt säkerhetsfrågor. Dessa data tyder på att säkerhetsrapportering i stor utsträckning negligeras i prövningar av interventioner inom psykisk ohälsa, kanske till och med i större utsträckning än inom andra medicinska områden. Denna brist hindrar bedömningen av förhållandet mellan risk och nytta vid ett rationellt beslutsfattande inom mentalvården. CONSORT-uttalandet utvidgades nyligen i ett försök att ge vägledning för en mer ändamålsenlig och omfattande rapportering av skadeverkningar i kliniska prövningar (4). Det återstår att se om rapporteringen av skador kommer att förbättras i framtiden.

EN JÄMFÖRELSE MELLAN STORA OCH SMÅ STUDIER

En fråga av betydande intresse är om det finns diskrepans mellan stora och små randomiserade kontrollerade prövningar vid frågor om psykisk ohälsa. Denna fråga har uppmärksamats inom flera andra medicinska områden (5). Vi tog nyligen upp den i samband med interventioner vid psykisk ohälsa (6). Vi gick igenom 161 systematiska Cochrane- och 254 DARE-översikter om interventioner i samband med psykisk ohälsa. Bland dessa identifierade vi 16 metaanalyser med åtminstone en "stor" randomiserad prövning med sampelstorlek > 800 och åtminstone en mindre prövning. Effektstorlekar beräknades särskilt för "stora" och mindre prövningar. Signifikant heterogenitet mellan studierna av olika storlek sågs i fem metaanalyser. Då man beräknade slumpmässiga effekter, skilde sig resultaten av stora och mindre prövningar utöver slumpen i 4 metaanalyser (25 pro-



FIGUR 1. Relativa förändringar av oddskvoterna för återfall i förhållande till den kumulativa sampelstorleken. Relativa förändringar större än 1 visar att nya studier lett till större oddskvoter i metaanalysen, dvs de kumulativa sammanfattande prövningsresultaten är mindre gynnsamma för interventionen under utvärdering ("experimentgruppen").

cent). I tre av dessa diskrepansfall (effekten av dagvård på intelligenskvot, utsättande av antidepressiva läkemedel, risperidon jämfört med typiska antipsykotika vid schizofreni) visade mindre prövningar större effekt än de stora prövningarna. I ett fall sågs det motsatta (olanzapin jämfört med typiska antipsykotika vid schizofreni). I fyra metaanalyser skilde sig effektstorleken mer än dubbelt mellan stora och mindre prövningar. Olika kvalitets- och designparametrar identifierades som potentiella förklaringar till vissa diskrepanser. Den första slutsatsen av denna undersökning är att stora prövningar är sällsynta vid psykisk ohälsa. Flera stora prövningar borde genomföras – och så länge de inte finns har metaanalyser en stor uppgift att fylla. Resultaten av stora prövningar överensstämmer oftast med resultaten av mindre studier, men diskrepanser

förekommer, så man måste vara försiktig när endast begränsade data finns att tillgå. Både stora och mindre prövningar borde granskas noggrant eftersom de utgör en obruten kedja av meningsfull evidens.

HUR MYCKET ÄNDRAS EVIDENSEN MED TIDEN?

Längs samma linjer undersökte vi också om säkerheten och uppskattningen av effekten och tolerabiliteten av interventioner vid psykisk ohälsa förändras med tiden, då flera prövningar i samma ämne tillkommer. Resultaten av metaanalyser kan och bör alltid uppdateras när nya data publiceras, och denna princip har accepterats av Cochrane-samarbetet åtminstone för alla Cochrane-översikter. Vi undersökte empiriskt 100 metaanalyser

(1 024 prövningar med 99 303 deltagare) där utfallet var död, återfall, misslyckande eller bortfall, med 5 prövningar publicerade under 3 olika år (7). Åtta metaanalyser nådde vid någon punkt formell statistisk signifikans ($p < 0,05$) men miste senare denna signifikans när data från flera prövningar samlades. I typiska fall upplöstes här stora effektstorlekar vid tidiga prövningar i och med ytterligare evidens. Med 500 randomiserade försökspersoner kunde förändringarna i oddskvoten senare bli mer än 1,5-faldiga i 95 procent av fallen. För utfallen död, återfall (Figur 1) och misslyckande var en minskning av effektstorleken något vanligare än en ökning av effektstorleken när fler data blev tillgängliga (157 mot 125, $p = 0,06$). Detta var tydligast vid jämförelser mellan läkemedelsbehandlingar och placebo (79 mot 51, $p = 0,009$). Utgående från dessa fynd slöt vi oss igen till

att evidens som baserar sig på ett litet antal deltagare i randomiserade prövningar bör tolkas försiktigt vid alla interventioner inom psykisk ohälsa. Speciellt kan en tidig effekt av läkemedelsterapi ibland överskattas.

SLUTSATSER

Sammantaget kan metaanalyser erbjuda ett sätt att få svar på specifika frågor av intresse genom att man ställer olika delar av tillgänglig evidens mot varandra. Vid sidan av tillämpningar som ger svar på enskilda frågor, är det lika eller kanske ännu mer intressant att använda dessa metoder för att få insikt i tillförlitligheten och osäkerheten i fråga om kliniska forskningsfynd i allmänhet. Att identifiera bristfälligheter kan bidra till bättre planerade studier och klarare kliniska svar i framtiden.

JOHN P.A. IOANNIDIS, MD

INSTITUTIONEN FÖR HYGIEN OCH EPIDEMIOLOGI, MEDICINSKA FAKULTETEN VID UNIVERSITETET I IOANNINA, IOANNINA 45110, GREKLAND

INSTITUTE FOR CLINICAL RESEARCH AND HEALTH POLICY STUDIES
TUFTS-NEW ENGLAND MEDICAL CENTER,
TUFTS UNIVERSITY SCHOOL OF MEDICINE,
BOSTON, USA

jioannid@cc.uoi.gr

REFERENSER

1. Lau J, Ioannidis JPA, Schmid CH. Summing up evidence: one answer is not always enough. *Lancet* 1998; 351:123-7.
2. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, Gotzsche PC, Lang T; CONSORT GROUP (Consolidated Standards of Reporting Trials). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*; 2001, 134:663-94.
3. Papanikolaou PN, Churchill R, Wahlbeck K, Ioannidis JP. Safety reporting in randomized trials of mental health interventions. *Am J Psychiatry*; 2004;161:1692-7.
4. Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, Moher D; CONSORT Group. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med*; 2004;141:781-8.
5. Cappelleri JC, Ioannidis JP, Schmid CH, de Ferranti SD, Aubert M, Chalmers TC, Lau J. Large trials vs meta-analysis of smaller trials: how do their results compare? *JAMA*; 1996;276:1332-8.
6. Contopoulos-Ioannidis DG, Gilbody SM, Trikalinos TA, Churchill R, Wahlbeck K, Ioannidis JP. Comparison of large versus smaller randomized trials for mental health-related interventions. *Am J Psychiatry*; 2005;162:578-84.
7. Trikalinos TA, Churchill R, Ferri M, Leucht S, Tuunainen A, Wahlbeck K, Ioannidis JP; EU-PSI project. Effect sizes in cumulative meta-analyses of mental health randomized trials evolved over time. *J Clin Epidemiol*; 2004;57:1124-30.