

# Låga p-värden är inte allt!

■ I Läkartidningen 34/2002 (sidorna 3302-5) skrev Adam Taube och Thomas Högberg om risken med missbruk av p-värden sedan lätthanterliga statistiska programpaket blivit tillgängliga.

Det är förvånande att så mycket uppmärksamhet fokuseras på p-värden. Har insikten om vad ett p-värde står för gått förlorat? Från att ha varit ett uttryck för sannolikhet ser det ut som om lågt p-värde numera står för det sanna. Men låt oss friska upp några basala begrepp.

## Samband

Finner vi ett samband mellan, låt oss säga ett förebyggande läkemedel och sjukdomsrisk, så finns sambandet i den undersökta gruppen. Det skall inte ifrågasättas. Det vi gärna vill veta är om sambandet också finns i den population som individerna kom ifrån. Ett av problemen vi då stöter på är risken att våra individer inte är representativa för hela populationen. Slumpen kan ha spelat oss ett spratt så att den grupp vi undersökte reagerade annorlunda på behandlingen än resten av populationen skulle ha gjort. P-värdet är ett mått på den risken och ingenting annat. Ett lågt p-värde, vanligen mindre än 0,05, brukar tolkas som att sambandet vi funnit också finns i hela populationen. I praktiken betyder det att vi accepterar att ta fel var tjugonde gång. (Jag vet att statistiskt skolade kolleger har en mer stringent definition men den här enkla varianten duger för det mesta). Gör vi många signifikanstester men låter gränsen för vad som uppfattas som statistiskt signifikant i vart och ett av testen ligga kvar vid  $p < 0,05$  så ökar naturligtvis risken att vi någon gång tar fel.

## Ett vanligt missbruk

Ett mycket vanligt missbruk av p-värden (även i fina familjer) är att man testar hurvida skillnader är signifikanta när man jämför deltagare i studier med ett bortfall, dvs de som skulle ha kunnat delta men vägrat, avbrutit behandlingen eller av någon annan orsak inte kunnat utvärderas. Det kan vara så att de som deltog i studien var något yngre än bortfallet. Med lite eftertanke förstår man att sådana skillnader är reella. Liksom ovan gäller det faktiska skillnader inom den undersökta gruppen, som i detta fall består av deltagarna och bortfallet. Om denna skillnad är av sådan art och storleksordning att den kan antas påverka den testade hypotesen så gör den det oavsett p-värde.

Multivariata analyser, där man samtidigt tar hänsyn till flera olika faktorer inverkan, föregås ibland av att man tes-



tar vilka faktorer som var för sig med en viss grad av statistisk signifikans korrelerar med utfallsvariabeln. Sedan väljer man att inkludera dem som korrelerade signifikant i sin multivariata analys och bortser från att andra faktorer som ändå mycket väl kan ha betydelse för studien. Man glömmer att »ej signifikant« inte är detsamma som »saknar betydelse«.

## P-värden används ofta felaktigt

Vad vill jag nu ha sagt med detta? Jo, att p-värden ofta används felaktigt och att de, när de används korrekt, är långt ifrån allt som har betydelse för värderingen av forskningsresultat. Inget p-värde i världen garanterar att ett samband funnet hos en grupp människor gäller för hela populationen. Även i stora undersökningar, där risken för felslut av rent slumpmässiga skäl är liten, går det ofta snett av andra orsaker.

## Många störfaktorer

I icke randomiserade studier förekommer ofta störfaktorer som måste beaktas vid resultatbetningen. Att justera för ålder och kön är vanligt. Därutöver finns ofta många kända, och dessvärre också okända prognostiska faktorer som behöver beaktas. Resultatets tillförlitlighet är beroende på i vilken utsträckning detta lyckas.

## Randomisering är inte hela lösningen

Randomisering används i kliniska prövningar för att aktiv behandlingsgrupp och kontrollgrupp skall bli lika. En randomisering medför dock inte automatiskt att de randomiserade grupperna liknar varandra i alla avseenden. Allt en korrekt randomisering tillför är att försökspersonens (patientens) tillstånd inte påverkar valet av behandlingsalternativ när patienten inkluderas i en prövning. Man undviker på så sätt »indication bias«.

Andra olikheter som kan uppstå får man ta hänsyn till vid resultatbearbetningen. Gäller det ett test av olika be-

handlingsalternativ är det viktigt att de patienter som ingår i prövningen liknar de patienter som senare kan bli föremål för behandling. I annat fall är det tvivelaktigt om resultat av prövningen är giltigt för vanliga patienter.

## Störd tolkning av resultaten

En störfaktor som sällan framhävs är vad som på engelska benämns »interpretation bias« [1]. Tolkningen av resultaten är inte oberoende av vem som tolkar och vilka medvetna eller omedvetna motiv som påverkar. Framför allt finns en risk att konsekvenserna av vad man funnit övervärderas [2]. Slutligen har vi »publication bias«. Det kan hos forskare eller finansierare finnas ett intresse att inte publicera vissa resultat. Det kan också vara så att vissa studier, framför allt med negativa resultat, kan vara svåra att få acceptans för i vetenskapliga tidskrifter.

Således, ett lågt p-värde garanterar på intet sätt att studieresultatet har allmän-giltighet. Minst lika viktigt är att uppmärksamma hur möjliga störfaktorer har hanterats, resultatens generaliserbarhet samt den kliniska relevansen.

## Gunnar Lindberg

*docent i allmänmedicin  
och epidemiolog vid NEPI*

*(Nätverk för läkemedelsepidemiologi)  
gunnar.lindberg@nepi.net*

## Referenser

1. Lindbaek M. Seeing what you want to see in randomised controlled trials. Meta-analyses may suffer from interpretation bias too. *BMJ* 2000;320:1720-3.
2. Werkö L, Hernborg A, Liedholm H, Melander A. Forskare som marknadsför sig – brist på omdöme eller ny tendens? *Läkartidningen* 2002;99:1962-3.

## Replik:

# Nej, just det!

■ »Låga p-värden är i inte allt!« utropar Gunnar Lindberg. Så sant, så sant – det var bl a därför som vi skrev vår artikel.

## Thomas Högberg

*docent, överläkare, kliniken  
för gynekologisk onkologi,  
Universitetssjukhuset, Linköping  
thomas.hogberg@ibk.liu.se*

## Adam Taube

*professor, statistiker,  
institutionen för informationsvetenskap  
(statistik), Uppsala universitet  
adam.taube@dis.uu.se*