

# Vad är egentligen en

Min upplevelse är att ordet trend ibland används märkligt av både statistiker och icke-statistiker. Vad menas när man säger att det är en "trend mot signifikanta resultat"? Och om man inte skall använda ordet trend i den här situationen hur skall det då användas?

Jag tror det flesta av oss är överens om att ordet "trend" oftast används (och då enligt min mening korrekt) för att beskriva en utveckling över tid och att det kan användas för att beskriva förändringar inom många vitt skilda områden som till exempel mode, kultur, utbildning. Kikar vi i SAOL så ges beskrivningen: "utvecklingsriktning, tendens, moderiktning", och visst är det så att vi ofta använder ordet trend för att beskriva samhällsförändringar och ofta kopplat till modetrender. Denna användning av trend är kristallklar och lämplig och här tror jag vi alla har samma förståelse av ordet "trend".

## Vad menas med "trend" i statistiska diskussioner?

Inom statistik verkar ordet "trend" ha fått en alldeles egen tilläggsbetydelse. Som diskuterat i tidigare artiklar i Pharma Industry så väljer vi oftast att säga att resultat är statistiskt signifikant om p-värdet  $<0,05$ , det vill säga om det är mindre än 5 procents chans att observera en så stor skillnad i behandlingseffekt som vi faktiskt gör, om sanningen är att det inte är någon skillnad mellan grupperna. Så vad säger vi då om p-värdet istället blir 0,06, 0,07 eller tom 0,09? Jo, många väljer då att säga att det är "en trend mot signifikant resultat". Vi vill ju SÅ GÄRNA ha statistisk signifikans och



# (statistisk) trend?

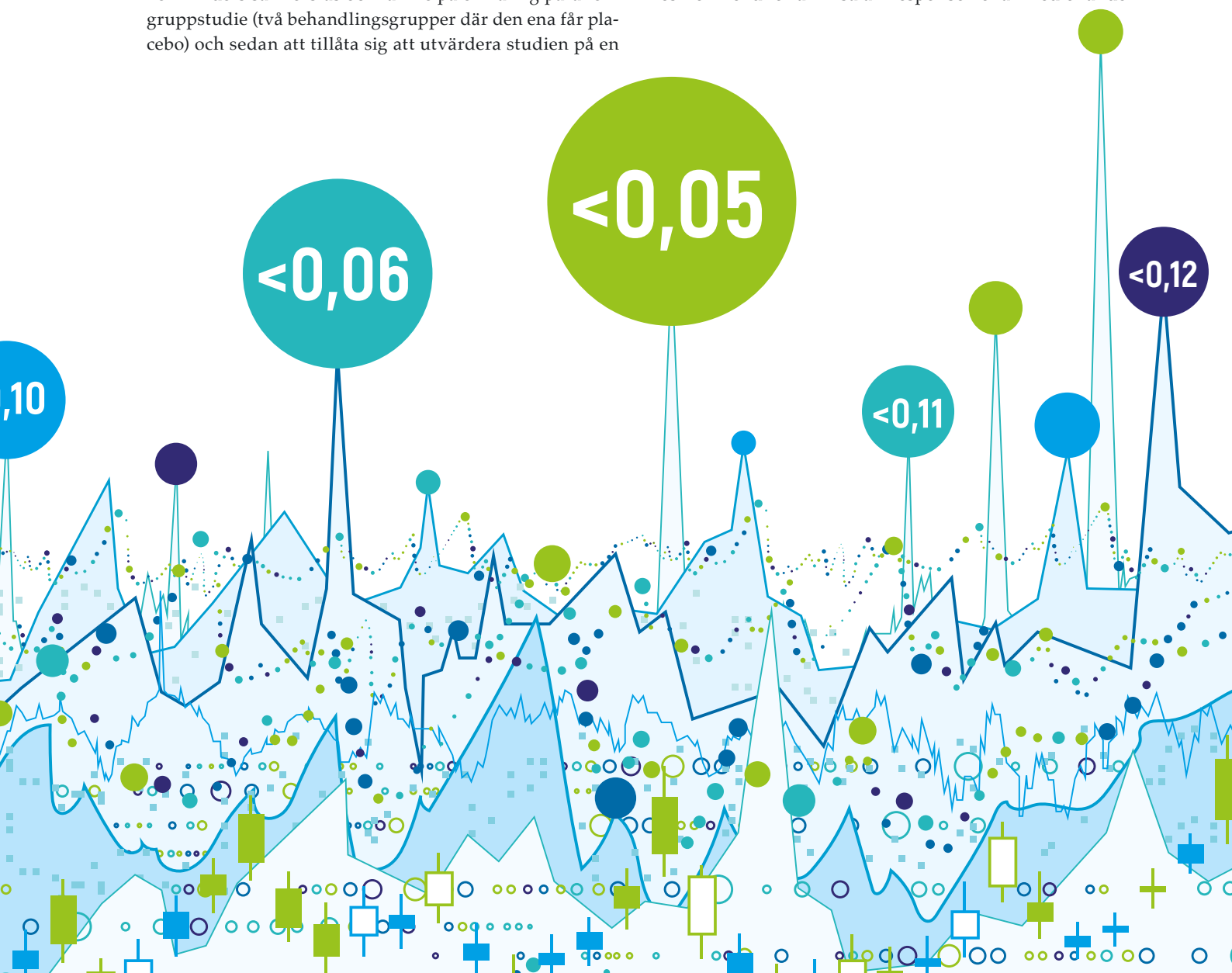
har vi inte det så kan det ju i alla fall omnämnas som en "trend mot signifikant resultat". Det man egentligen menar är att man nådde nästan ända fram till den magiska gränsen 0,05. Men nära och nästan känns inte så övertygande, så för att få det att låta lite bättre slänger vi in ordet "trend", som om vi liksom är på väg även om p-värdet är 0,08 JUST NU.

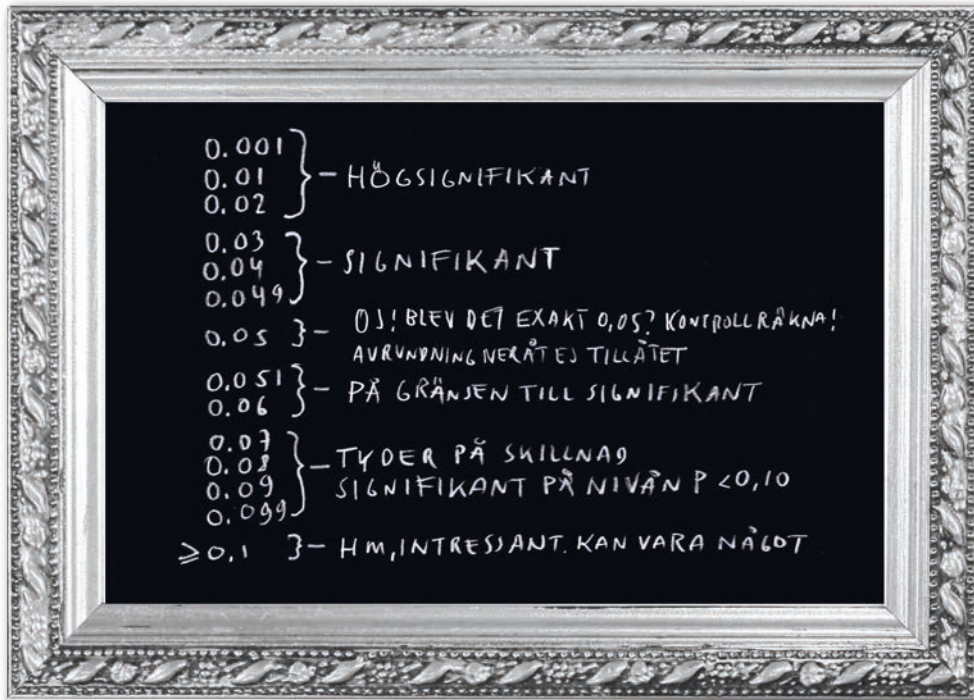
Så varför är detta ett problem? Varför kan vi inte få omnämna nästan-signifikanta resultat som en "trend mot statistiskt signifikanta resultat"? Jo, för att det kan bli så fel om mottagaren sitter och tänker på en annan betydelse av ordet trend. Nyligen tillbringade jag ett par timmar i ett möte med min gode vän Anders där vi diskuterade studiedesign inför en fas II-studie. Frågan gällde om en studiedesign inriktad mot trendanalys kunde minska studiestorleken. Anders satt förstås och tänkte på en vanlig parallellgruppstudie (två behandlingsgrupper där den ena får placebo) och sedan att tillåta sig att utvärdera studien på en

högre signifikansnivå, till exempel  $p=0,10$ . Jag satt och tänkte på att inkludera flera dosnivåer för att kunna utnyttja en eventuell linjär trend för att på så sätt lättare kunna visa effekter. Efter knappt två timmar landade vi i samma förståelse av begreppet trend.

## Trend är bara ett bra ord när vi faktisk pratar om trender

I kliniska studier finns det flera situationer där det faktiskt är intressant att använda statistiska test för trend. Det kan till exempel gälla när vi utvärderar en utveckling över tid, eller nästan ännu vanligare när man jämför respons för flera dosnivåer. Har vi en studie med flera olika dosnivåer så kan det vara lättare att använda till exempel ett statistiskt test för trend för att visa att responsen ökar med ökande





Förslag till tolkning av p-värden. Bild: Anders Gunér.

dos och att tilläggsinformationen vi får av att använda flera olika doser bidrar till att vi lättare ser effekter. Det finns alltså ett tydligt användningsområde för trendtest inom statistiken. Gemensamt för dessa är att man behöver data från flera mätpunkter (över tid eller för flera dosnivåer) och att man utnyttjar informationen som ligger i att mätpunkterna är ordnade efter dos eller tid.

### Borderline significant, on the edge of significant, numerically superior\*...

Vad betyder då alla dessa formuleringar? Det är ju inte bara ordet trend som känns svårtolkat. I de allra flesta fall är detta bara varianter på temat "vi skulle så gärna önska att p-värdet var lite lägre" och man förstärker budskapet genom att använda starkare ord. Begreppet "numerically superior" tycker jag är speciellt charmigt. Det lärde jag mig under min tid som Clinical Research Associate i ett stort läkemedelsbolag där uttrycket användes flitigt för att antyda att resultaten för prövningsläkemedlet var bättre än för referensbehandlingen: endpointen i studien var en score (poängsumma) och prövningsläkemedlet hade ca 1 poäng högre score än referensläkemedlet. Helt negligerbart effektmässigt och inte ens i närheten av att kunna kallas "trend mot signifikant resultat" som ändå kräver ett någorlunda lågt p-värde.

### Hur nyanserar vi ett icke-signifikant resultat?

Eftersom tolkning av p-värden är krångligt nog så vi bör vi använda så enkla ord som möjligt. Om vi absolut behöver beskrivande ord, varför då inte bara skriva "nästan signifikant" eller "signifikant på 10 procentsnivån (om  $p < 0,1$ )"? Om man vill nyansera och förstärka ytterligare så är det ju inte bara p-värdet som är intressant. P-värdet är bara ett mått på om de resultat vi ser kan ha orsakats av en

slump eller om det sannolikt finns en sann underliggande skillnad i behandlingseffekt. Störst fokus bör vi lägga på skattning av effekt, punkttestimat och konfidensintervall.  $P=0,09$  i en studie betyder inte nödvändigtvis vara samma sak som  $p=0,09$  i en annan studie. I en mindre studie har vi små möjligheter att nå statistisk signifikans på grund av få patienter och därmed låg statistisk styrka. Den effektstorlek man observerar kan ändå vara intressant och bidra till att man väljer att göra en större konfirmerande studie.

### Kan man avrunda 0,051 till 0,05?

Jag avrundar med en anekdot från min tid som mer praktisk statistiker. Jag fick ett telefonsamtal från en känd läkare på ett större sjukhus i Sverige som arbetade med att sammanställa sina resultat från en klinisk studie. Han hade räknat sig fram till det mest avskydda resultatet av alla:  $p=0,051$ . Ungefär som 4:e plats i OS eller VM. Han argumenterade med att enligt normala avrundningsregler borde detta kunna avrundas till 0,05 och han ville ha okej på detta. Vad jag svarade? Jag svarade att omständigheten att han behövde ringa mig för att fråga om han kunde avrunda ner till 0,05, borde redan gett honom det svar han behövde. Jag tror alla är överens om att fiffla med siffrorna inte är okej. Att uppfinna nya adjektiv för att göra resultaten mer intressanta verkar däremot vara *comme il faut*.

\*Ursäktla engelskan, men nästan all vetenskaplig litteratur är på engelska så det känns mer meningsfullt att inte översätta.



**ANNA TÖRNER**  
statistiker och verkställande direktör  
Scandinavian Development Services