

Rangordning i vården leder till felaktiga slutsatser

Rangordning har blivit en populär metod för jämförelser av olika enheters resultat inom hälso- och sjukvården. Men risken är stor att denna metod leder till förhastade och ovetenskapliga slutsatser, med negativa effekter för förbättringsarbetet. Rangordning är således en dålig metod för att jämföra kvalitet och effektivitet, skriver fyra forskare från Lund och föreslår ett bättre alternativ.

Sveriges Kommuner och Landsting (SKL) och Socialstyrelsen publicerar årligen rapporten »Öppna jämförelser av hälso- och sjukvårdens kvalitet och effektivitet« [1] för att ge »en heltäckande bild av verksamheternas kvalitet och effektivitet« [2]. Jämförelserna har hittills bestått i rangordning av landstingsverksamhet, men strategin är att rapportera även på sjukhus- och klinikknivå [3].

Rangordningsjämförelser förekommer också i andra länder. I Storbritannien har det t ex under många år publicerats »league tables« för skolor och sjukhus. Detta har lett till att rangordningsjämförelser identifierats som ett allvarligt metodologiskt problem [4]. Ledande statistiker har rekommenderat att redovisningarna bör upphöra [5].

Vi vill varna för förhastade slutsatser av rangordning och uppmuntra till stor försiktighet med tolkning av resultaten.

Vad som är det sanna genomsnittliga resultatet av en viss behandling vid ett visst sjukhus kan bara bestämmas för definierade grupper av redan behandla-

de patienter. Sådana resultat avspeglar emellertid endast historiska förhållanden, påverkade av tillfälligheter, och kan inte utan vidare användas för jämförelser av framtida behandlingsresultat.

Det observerbara genomsnittliga resultatet av en behandling vid ett sjukhus är inte konstant. Olika urval av patienter som får samma behandling har olika genomsnittresultat. Denna sjukhus-specifika variabilitet måste beaktas för att jämförelser mellan sjukhus ska vara meningsfulla.

Anta exempelvis att ett gammalt välrenommerat sjukhus under 2007 redovisat att 14 800 patienter behandlats för en viss allvarlig sjukdom och att 3 700 överlevt (25 procent överlevnad). Ett nystartat sjukhus för alternativ medicin redovisar samtidigt att en patient framgångsrikt behandlats för samma sjukdom och överlevt hela året (100 procent överlevnad).

Bortser man från urvalsosäkerheten

kan man konstatera att det nya sjukhuset var betydligt mera framgångsrikt än det gamla eftersom överlevnaden var 100 procent jämfört med 25 procent. Då måste man emellertid också acceptera att resultatet vid det nya sjukhuset enbart gäller den redan behandlade patienten och

saknar relevans för andra, framtida, patienter.

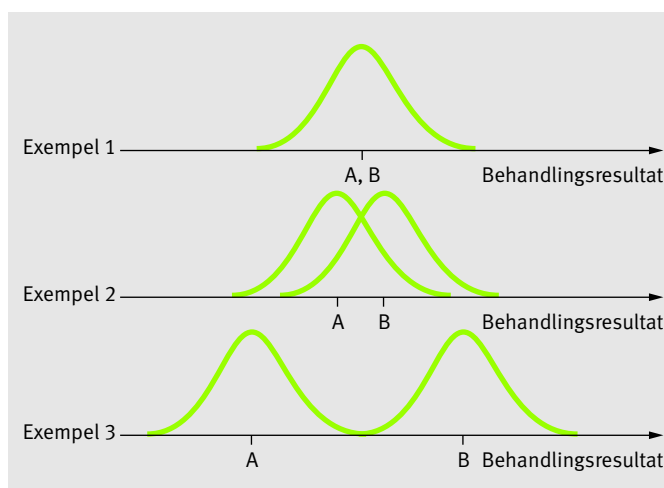
Betraktar man i stället sjukhusens behandlingar under 2007 som slumpurval och använder observerade genomsnittsvärden som skattningar av sjukhusens totala genomsnittresultat, blir jämförelserna intressanta ur ett allmänt patientperspektiv och kan tolkas med utgångspunkt från kvalitet och effektivitet.

Urvalsosäkerheten måste då naturligtvis beräknas, och en felmarginal anges för varje skattning. Detta görs lämpligen i form av ett 95-procentigt konfidensintervall, konstruerat så att det med 95-procentig säkerhet omfattar det skattade värdet. Felmarginalerna i exemplets överlevnadstal, 25 procent och 100 procent, blir ungefär 24–26 procent respektive 0–100 procent.

Någon meningsfull effektivitets- och kvalitetsorienterad jämförelse med det nya sjukhuset kan alltså inte göras.

Ju större osäkerhet och homogenitet som kännetecknar observerade genomsnittliga behandlingsresultat, desto mindre meningsfullt är det att rangordna dem, se Figur 1.

Urvalsosäkerheten för två sjukhus A och B symboliseras med normalfördelningar. Varje fördelning beskriver den uppsättning genomsnittliga behandlingsresultat som kan observeras i olika



Figur 1. Behandlingsresultat och urvalsosäkerhet.

JONAS RANSTAM

driftschef, NKO, ortopediska kliniken
jr@nko.se

PHILIPPE WAGNER

statistiker, NKO

OTTO ROBERTSSON

överläkare, ortopediska kliniken

LARS LIDGREN

professor, ortopediska kliniken; samtliga vid Universitetssjukhuset i Lund



Foto: Lennart Hyse / Scampix

Rangordning – en metod som kanske lämpar sig bättre för idrotten än inom sjukvården?

slumpurval från respektive sjukhus.

I extremfallet att sjukhusen har identiska totala genomsnittliga behandlingsresultat (Exempel 1) kommer all rangordning att bli missvisande och enbart avspeglar urvalsskillnader.

I det andra extremfallet är rangordningen tillförlitlig (Exempel 3). Den kombinerade urvalsosäkerheten är då mindre än skillnaderna i de totala genomsnittliga behandlingsresultaten.

I »Öppna jämförelser« redovisad variabilitet i observerade behandlingsresultat och deras konfidensintervall tyder inte på att genomförd rangordning är tillförlitlig.

Ett annat problem för rangordning är att registerdata inte alltid är korrekta. Konsekvenserna av detta är förmodligen i många fall försumbara jämfört med urvalsosäkerheten, men låt oss för resonemangets skull göra det tvivelaktiga antagandet att observerade genomsnittliga behandlingsresultats rangordning är meningsfull i sig.

En nyligen genomförd simuleringsstudie [6], baserad på av Höftprotesregistret publicerade data om reoperation inom 2 år vid 77 sjukhus [7], visar att redan vid 2–3 procent oregistrerade operationer får hälften av sjukhusen fel rang, och felet kan vara avsevärt – för vissa sjukhus uppåt 20 placeringar.

Svenska kvalitetsregister har i bästa fall en kompletthet av ungefär 96–97 procent [8, 9].

Det första nationella kvalitetsregistret, Svenska Knäprotesregistret, har sedan starten 1975 redovisat [10] klart minskad risk för omoperation. Reoperationsrisken har i Sverige under de senaste två decennierna halverats och är nu lägst i världen. Skillnader i klinikresultat har också minskat avsevärt. Detta gäller även höftproteskirurgin.

Förbättringen har åstadkommit i en löpande dialog med inrapporterande enheter, där brister i patienturval, protes och kirurgisk teknik redovisats, men också med årlig återrapportering av kliniskspecifika resultat, utan att rangordna sjukhus.

En direkt konsekvens av minskade sjukhusskillnader är att rangordning fungerar allt sämre som förbättringsinstrument. Nya och bättre metoder att skatta sjukhusspecifika behandlingsresultat finns och har tagits i anspråk [11] men förändrar inget i sak. Som instrument att förbättra sjukvården riskerar rangordning i stigande grad att leda till »förbättringsförsämring« genom att dåliga förebilder framställs som goda och vice versa.

Urvalsosäkerheten i »Öppna jämförelser« rangordningar bör naturligtvis redovisas öppet, t ex med konfidensintervall för rangtal [12].

Ännu bättre är att helt överge rangordning och basera jämförelser på både observerat resultat och på resultatets säkerhet, exempelvis med kategorierna:

1) säkerställt gott resultat, 2) säkerställt dåligt resultat och 3) osäkert resultat. De tre kategorierna kan symboliseras med stoppljus eller varför inte som »smileys«.

Även jämförelser som tar hänsyn till urvalsosäkerhet blir emellertid med systematiska skillnader i patienters kön, ålder, sjuklighetsgrad, komorbiditet, etc missvisande. Rättvisa jämförelser kräver korrektion för obalans i prognostiska faktorer bland patienterna, ett omfattande arbete.

■ *Potentiella bindningar eller jävsförhållanden: Inga uppgivna.*

REFERENSER

1. Sveriges Kommuner och Landsting, Socialstyrelsen. Öppna jämförelser av hälso- och sjukvårdens kvalitet och effektivitet. Jämförelser mellan landsting 2006. Lindsberg: Bergslagens grafiska AB, juni 2006.
2. Sveriges Kommuner och Landsting. Öppna jämförelser – hälso- och sjukvård, grundskola och äldreomsorg. Cited 26 July 2008, available from: <http://www.skl.se/artikel.asp?C=6089&A=46674>
3. Sveriges Kommuner och Landsting, Socialstyrelsen. Nationell strategi för »öppna jämförelser« av hälso- och sjukvårdens effektivitet och kvalitet. Cited 26 July 2008, available from <http://www.skl.se/artikel.asp?A=20272&C=4809>
4. Goldstein H, Spiegelhalter D. League tables and their limitations: statistical issues in comparing institutional performance. *J Roy Stat Soc A*. 1996; 159:385–443.
5. Goldstein H, Leckie G. School league tables: what can they really tell us? *Significance*. 2008;5:67–69.
6. Ranstam J, Wagner P, Robertsson O, Lidgren L. Healthcare quality registers: outcome-oriented ranking of hospitals is unreliable. *J Bone Joint Surg Br* (in press).
7. Swedish Hip Arthroplasty Register. Annual Report 2006. Gothenburg: Department of Orthopaedics, Sahlgrenska University Hospital; 2007.
8. Robertsson O, Dunbar MJ, Knutson K, Lewold L, Lidgren L. Validation of the Swedish Knee Arthroplasty Register – A postal survey regarding 30,376 knees operated on between 1975–1995. *Acta Orthop Scand*. 1999;70(5):467–72.
9. Troëng T. Swedvasc's numbers are reliable. *Läkartidningen*. 2008;105:557.
10. Svenska Knäprotesregistret. Årsrapport 2007. Ortopediska kliniken, Universitetssjukhuset i Lund. Cited 26 Juli 2008, available from http://www.knee.nko.se/online/uploadedFiles/100_SKAR2007_Sv1.2.pdf
11. Robertsson O, Ranstam J, Lidgren L. Variation in outcome and ranking of hospitals: an analysis from the Swedish knee arthroplasty register. *Acta Orthop*. 2006;77:487–93.
12. Marshall EC, Spiegelhalter DJ. Reliability of league tables of in vitro fertilisation clinics: retrospective analysis of live birth rates. *BMJ*. 1998;316:1701–5.

Fler debattinlägg på sidan 2343.