

Anders Ahlbom, professor, Institutet för miljömedicin, Karolinska institutet, Stockholm, samt Samhällsmedicin, Stockholms läns landsting (anders.ahlbom@imm.ki.se)

Lars Alfredsson, professor, Institutet för miljömedicin, Karolinska institutet, Stockholm, samt Samhällsmedicin, Stockholms läns landsting

Tomas Andersson, statistiker, Samhällsmedicin, Stockholms läns landsting

Johan Hallqvist, docent, överläkare, Samhällsmedicin, Stockholms läns landsting, samt institutionen för folkhälsovetenskap, Karolinska institutet, Stockholm

Ulf Strömberg, professor, Yrkes- och miljömedicin, Lunds universitet; samtliga för Epinet

Fler nivåer, färre skillnader

Aktuell analys av sjukhusens vårdresultat kanske ger missvisande bild

I en artikel i detta nummer av Läkartidningen ger sig Merlo och medarbetare in i diskussionen om skillnader mellan olika sjukhus vad gäller dödlighet efter hjärtinfarkt. Deras slutsats är att de skillnader som finns är mycket små och att för en patient som byter sjukhus till ett med högre risk skulle dödsrisken »öka med i median« omkring 10 procent. Samtidigt kan man ur deras figurer läsa ut att det sjukhus med högst risk har en dödlighet som är mellan tre och fyra gånger så hög som det sjukhus med lägst risk (Figur 1 A och 2 A, sidan 22 i detta nummer); i den följande figuren har detta genom diverse åtgärder reducerats till omkring 50 procent (Figur 1 B och 2 B). Merparten av den ursprungliga 3–4-faldiga skillnaden har försvunnit genom att författarna använt en statistisk analysmodell som blir allt mer populär och som brukar kallas flernivåmodell.

Det finns många situationer där data är hierarkiskt uppbyggda och där sådana modeller kan komma ifråga. I det aktuella exemplet är patienter, sjukhus och landsting tre nivåer. I tvillingforskning är enskilda tvillingar en nivå och tvillingpar en annan. I metaanalys är individerna i en enskild studie en nivå och de olika studierna en annan nivå. Skälet till att sådana situationer kan påkalla särskilda analysmetoder är att det kan finnas ett statistiskt beroende mellan de individer som ingår i de olika grupperingar som bildas. Patienter på ett sjukhus kan skilja sig från patienterna på ett annat, och tvillingar i ett par har mer gemensamt med varandra än med tvillingar i andra par. Individerna i en studie i en metaanalys kan skilja sig från individerna i en annan studie.

Det kan följaktligen finnas underliggande skillnader i risk mellan sådana grupperingar. Ibland är dessa skillnader själva huvudsyftet med en studie, som t ex vid analys av sjukhuskillnader i död efter hjärtinfarkt. Det gäller också inom viktiga delar av social epidemiologi där sjukdomsrisker analyseras på individnivå, men också i en social eller geografisk så kallad kontext som alltså bildar en högre nivå. I andra situationer är dessa skillnader bara till besvär men måste ändå hanteras på ett adekvat sätt. Detta gäller som regel i metaanalyser; även om studieskillnader kan vara av stort intresse vore det som regel en fördel om de inte fanns.

Förment lättolkade parametrar kan missförstås eller feltolkas

När man talar om flernivåmodeller eller hierarkiska modeller menar man som regel modeller med inslag av »random effects«. Med en sådan effekt avses att inte bara utfallet för en enskild individ eller patient är styrt av en slumpmässig pro-

Sammanfattat

Det finns många situationer där data är hierarkiskt uppbyggda och där flernivåanalyser kan komma ifråga. Metoderna måste dock anpassas till frågeställningen, och modellerna måste kontrolleras mot data.

Se även artikeln på sidan 20 i detta nummer.

cess, utan också att den underliggande risken är slumpmässigt bestämd. Detta blir aktuellt vid klusterurval, till exempel av skolklasser där det kan finnas en klassnivå som alltså bestäms av vilka klasser som ingår i urvalet och dessutom en elevnivå.

För analyserna av sjukhusdödlighet innebär detta att om antalet avlidna hjärtinfarktpatienter vid ett sjukhus antas följa en binomialfördelning följer också binomialfördelningens parameter, »p«, en sannolikhetsfördelning; ofta antar man att oddskvoten följer en logaritmisk normalfördelning. Det är vanligt vid flernivåmodeller att man blandar fixa effekter och »random effects«, till exempel genom att använda vanliga fixa effekter på individnivå och »random effects« på högre nivåer. Man talar då om »mixed effects models« eller mixade modeller.

Ett problem med flernivåanalyserna är att resultaten levereras med diverse förment lättolkade parametrar som dock ofta missförstås eller feltolkas.

Traditionellt redovisas analyser av flernivåmodeller med en total varians som sedan delas upp i komponenter som hänförs till de olika nivåerna. Det går då att se hur stor del av variansen som kan hänföras till t ex patientnivån eller sjukhusnivån.

Detta är dock en förledande enkel resultatredovisning. Variansen i sig går inte att överföra till någon tolkningsbar skala, och samma varians kan rymma både mycket små och mycket stora sjukhuskillnader. Merlo och medarbetare anger t ex att variansen mellan landsting för kvinnor är 0,001 och mellan sjukhus 0,014. Det är inte lätt att förstå vad detta betyder annat än att talen är små. Trots det så uppvisar alltså de data som modellen bygger på en skillnad i dödsrisk som är

Det är av stor vikt att modellernas giltighet kontrolleras gentemot empiriska data. Det är också av största vikt att metoderna anpassas efter den aktuella frågeställningen.

uppmot fyra gånger högre när sjukhuset med den högsta respektive lägsta risken jämförs.

Mer relevant tolka medianoddskvot utifrån sjukhusnivå

Att dessa variansmått inte är särskilt relevanta för de flesta epidemiologiska frågeställningar har de statistiker uppmärksammat som utvecklar dessa metoder, och man har därför försökt översätta resultaten till mer tolkningsbara mått på skalor som används inom epidemiologin [1].

Merlo och medarbetare arbetar med logistisk regression, och det mått som rekommenderas i detta sammanhang är medianoddskvot. De påstår att detta är ett enkelt och lättolokat mått och att resultaten direkt kan jämföras med t ex en oddskvot som beskriver hur förekomst av diabetes påverkar dödligheten efter hjärtinfarkt. Men det är en sanning med modifikation: Om sjukhusen som ingår i analysen kombineras parvis på alla möjliga sätt och en oddskvot bildas för varje par, alltid med den högre dödligheten i täljaren erhålls ett antal oddskvoter. Dessa oddskvoter har en fördelning, och medianoddskvoten är medianen i denna fördelning. I praktiken beräknas dock värdet på medianoddskvoten utifrån antagandet om med vilken sannolikhetsfördelning som sjukhusdödligheten varierar mellan sjukhusen.

För kvinnor blir medianoddskvoten 1,12 och för män 1,10; detta beskrivs för läsaren som att om en kvinna kunde söka vård vid ett annat sjukhus med högre dödlighet skulle hennes dödlighet öka med i median 12 procent. Mer relevant än att tolka medianoddskvot utifrån en patients perspektiv är att tolka den på sjukhusnivå: en medianoddskvot på 1,12 innebär att oförklarade skillnader i kvinnliga patients dödlighet är mindre än 12 procent i hälften av alla parvisa jämförelser mellan svenska sjukhus.

Medianvärdet säger dock inte allt om de oförklarade, kvarvarande skillnader i dödlighet som finns på sjukhusnivå. Det är också viktigt att, som Merlo och medarbetare gör, beskriva skillnaden mellan sjukhus i figurer.

Håller modellen för att beskriva verkligheten?

I ett samhällsperspektiv är det också dessa figurer som har den väsentliga informationen. Däremot kan man undra om det är Figur 1 A och 2 A med 400 procents skillnad eller Figur 1 B och 2 B med 50 procents skillnad mellan sjukhus med högsta respektive lägsta risk, som har mest relevant information. Figur 1 A och 2 A är baserade på faktiska dödsrisker, och Figur 1 B och 2 B är baserade på oddskvoter med justering för fler variabler och med flernivåanalys med en mixad modell.

Figurerna skiljer sig alltså på mer än ett sätt. Värdet på medianoddskvoten är direkt relaterat till fördelningen i Figur 1 B och 2 B, och det är ett sätt att sammanfatta variationen mellan sjukhusen. Det är dock viktigt att komma ihåg att värdet på medianoddskvoten är baserat på ett modellantagande om hur fördelningen på sjukhusnivå ser ut. I det här fallet har man an-

II Fakta

Nätverket Epinet

Epinet är ett nätverk av svenska epidemiologer som träffas regelbundet för att diskutera metodfrågor av särskilt intresse.

Vid ett möte diskuterades metodproblem och tolkningsfrågor vid flernivåanalyser.

tagit att oddskvoterna följer en logaritmisk normalfördelning. Det antagandet måste kontrolleras, och det är bara om de faktiska effekterna följer den logaritmiska normalfördelningen som det är rimligt att använda den för att beskriva variationerna på sjukhusnivå.

Den stora skillnaden mellan Figur 1 A/2 A och Figur 1 B/2 B är att de mest avvikande värdena uppåt och nedåt har försvunnit i Figur 1 B/2 B. Tyvärr framgår det inte vad detta beror på, men möjligen är förklaringen att man genom antagandet tvingat in data i en modell som drastiskt jämnar ut fördelningen. Om detta är bra eller dåligt beror på om variationerna är slumpmässiga effekter eller en följd av reella sjukhusskillnader, vilket vi förstas inte vet. Det förefaller ändå rimligt att med 3- till 4-faldiga skillnader titta något närmare på sjukhusen med den högsta respektive lägsta risken och att följa utvecklingen över tid, snarare än att konstatera att de är förenliga med slumpvariationer.

Framtida erfarenheter får utvisa tillämpbarheten

Sammanfattningsvis kan man konstatera att modeller med »random effects« används alltmer i flernivåsituationer, och vi delar Merlos och medarbetares uppfattning att detta är bra, även om framtida erfarenheter får visa hur stor tillämpbarheten är i olika situationer och hur stora de praktiska konsekvenserna är.

Det är av stor vikt att modellernas giltighet kontrolleras gentemot empiriska data. Det är också av största vikt att metoderna anpassas efter den aktuella frågeställningen. När det gäller det aktuella exemplet är vi inte övertygade om att den mixade modell som Merlo och medarbetare använder är väl förenlig med data, eftersom modellen drastiskt jämnar ut skillnaderna i dödsrisk.

*

Potentiella bindningar eller jävsförhållanden: Inga uppgivna.

Referenser

1. Larsen K, Petersen JH, Budtz-Jorgensen E, Engdahl L. Interpreting parameters in the logistic regression model with random effects. *Biometrics* 2000;56:909-14.