

Variasjon i dødelighet mellom sykehus må tas alvorlig

REDAKSJONELT

GULBRANDSEN P

HELTEFs rapport er et godt utgangspunkt for studier av egen virksomhet

Stiftelse for helsetjenesteforskning (HELTEF) lanserte i mars sin rapport om variasjon i dødelighet for fem utvalgte diagnoserelaterte grupper (DRG) ved norske sykehus (1). Som man kunne vente, har rapporten fått stor oppmerksomhet, og det må sies å ha vært temmelig naivt av HELTEF å tro at politikere og presse ville slå seg til ro med at det enkelte sykehus ikke ville bli eksponert. I studien forsvares anonymiteten med at datagrunnlaget er for usikkert til at dårlige resultater vil kunne rettsfølges. Det har forfatterne helt rett i, men det er ikke noe argument for å holde tilbake opplysninger om offentlig virksomhet. Både sykehuseiere og leger må nå innse at åpenhet om behandlingsresultater er kommet for å bli. Bristol-saken viste vei (2). Den store utfordringen er å forklare publikum de begrensninger en slik undersøkelse har, og redusere frykten mange vil føle overfor innleggelse i sykehus med relativt dårligere resultater enn andre. Det er klart folk blir forvirret når forskjellen i absolutt (ukorrigert for alder, kjønn mv.) dødelighet for lårhals- og lårbeinsbrudd mellom det “dårligste” og det “beste” sykehuset er på 0,04%, mens dette er fremstilt som en relativ forskjell på 100% (3).

Men det er også en utfordring å få sykehuseiere og leger til å ta resultatene alvorlig, og ikke feste seg ved alle mulige feilkilder. Rapporten gir en klar redegjørelse for metodesvakheter, og seks vesentlige forklaringer på den store variasjonen i dødelighet diskuteres.

For det første er det ikke kontrollert for ulik sykелighet i sykehusenes nedslagsfelt. Man kan f.eks. tenke seg at befolkningen i Sogn og Fjordane generelt er så mye sterkere enn andre at de bedre tåler et hjerteinfarkt. En mulig forklaring på at større sykehus generelt kommer noe dårlig ut kan være at de sykeste pasientene ikke legges inn i lokalsykehus. En tredje feilkilde kan være at man har ulike terskler for sykehusinnleggelse i ulike deler av landet. Dette gjelder nok især pasienter med hjertesvikt, lungebetennelse eller pleuritt, i mindre grad dem med hjerteinfarkt og lårhals- eller lårbeinsbrudd. Videre kan ulik prehospital dødelighet betinget i distanse til sykehus og kvalitet på transporttjenesten være en feilkilde, likeså mulige forskjeller i diagnoseklassifisering.

Når så mange sammenlikninger gjøres som i denne rapporten, må dessuten tilfeldig variasjon forventes. Selv ved en randomisering av innleggelsene til 60 helt like sykehus ville man vente signifikant avvikende resultater for tre sykehus per diagnose man studerer.

Men antallet avvik i HELTEFs undersøkelse var fem ganger større enn dette, og kan ikke forklares med tilfeldigheter alene. Derimot burde kanskje Bonferroni-korreksjon vært

anvendt for å korrigere for antall gjennomførte analyser. Forfatterne begrunner fraværet av denne korreksjonen med at 95% konfidensintervall for resultatene allerede representerer en meget stor feilmargin med lav sensitivitet for å identifisere sykehus med avvik. Fra et pasientsynspunkt er så lav sensitivitet neppe ønskelig, hevder de, og man valgte derfor pragmatisk en presentasjon av enkeltresultater uten korreksjon. Dette skjønn kan diskuteres. Hvis datagrunnlaget er svært usikkert, bør marginene antakelig være store.

Studien søker også etter mulig samvariasjon mellom sykehusets aktivitetsnivå og dødelighet, uten å kunne påvise overbevisende sammenhenger. Denne analysen bygger på en teoretisk modell som presenteres altfor kortfattet i rapporten. Analysene dokumenteres med spredningsdiagram med regresjonslinje. Bruken av en slik linje er ofte misvisende og bør unngås i korrelasjonsanalyser (4).

Rangering av behandlingsresultater basert på registerdata er beheftet med stor usikkerhet (5). Forfatterne underslår ikke dette. De redegjør for at rapporten først og fremst må brukes av det enkelte sykehus for å stille spørsmål ved egen virksomhet. Rådet er at dersom man kommer dårlig ut for en av diagnosegruppene, bør dette føre til en analyse av hva som kan være forklaringen. Det kan godt være at årsaken er forhold sykehuset selv ikke er herre over, men det er ikke et svar man kan slå seg til ro med uten en grundig vurdering. Og skulle årsaken være avvikende diagnoseklassifisering, og ikke et reelt dårligere behandlingsresultat, burde også det føre til endring. Alle er tjent med at oppbyggingen av databaser gjøres med så lite variasjon i kvaliteten på dataene som legges inn, at det er mulig å gjøre reelle sammenlikninger, ikke minst over tid.

Analysene bygger på dødsfall under sykehusoppholdet. Forfatterne anbefaler at man registrerer dødsfall inntil 30 dager, ev. også 90 dager, etter utskrivning. De frykter at noen vil ønske å manipulere sine resultater ved å skrive ut døende pasienter, og derved forhindre at dødsfallene vil bli registrert i analyser som den foreliggende. Så absurd en slik tanke enn måtte synes, kan vi dessverre ikke se bort fra slike konsekvenser. Datamanipulasjon forekommer i kjølvannet av at dødelighetsdata publiseres (6).

Denne type rapporter vil vi komme til å se flere av. De representerer et viktig verktøy for kvalitetsforbedring i helsevesenet, men det er nødvendig å understreke at de absolutte forskjeller i dødelighet mellom sykehusene ikke er dramatiske. Trolig reflekterer resultatene vel så mye feilkildene som reelt store forskjeller i behandlingskvalitet. Likevel er det ikke gunstig at man, ved å ta rapporten "med stor ro" eller betvile undersøkelsens kvalitet, gir publikum et signal om at man er tilfreds. Selv om denne reaksjonen kan være en konsekvens av innsikt i undersøkelsens svakheter, er det ikke slik man vil bli oppfattet i befolkningen. Helsevesenets holdning må være offensiv: Disse tallene vil vi studere nøye!

Pål Gulbrandsen

Pål Gulbrandsen (f. 1955) er medisinsk redaktør i Tidsskriftet.

LITTERATUR:

1. Guldvog B, Kopjar B. In-hospital mortality rates in Norway 1994-97. Rapport nr. 2/1999. Oslo: HELTEF, 1999.
2. Vonen B, Larsen M-L. Vi må ta Bristol-saken på alvor. Tidsskr Nor Lægeforen 1998; 118: 3174.
3. Frich L. Dødelighet ved sykehus til besvær. Aftenposten (morgenutgave) 29.3.1999.
4. Porter AMW. Misuse of correlation and regression in three medical journals. J R Soc Med 1999; 92: 123-8.
5. Marshall EC, Spiegelhalter DJ. Reliability of league tables of in vitro fertilisation clinics: retrospective analysis of live birth rates. BMJ 1998; 316: 1701-4.
6. Schneider EC, Epstein AM. Influence of cardiac surgery performance reports on referral practices and access to care. N Engl J Med 1996; 335: 251-6.

Publisert: 17. oktober 2018. Tidsskr Nor Legeforen. DOI:
© Tidsskrift for Den norske legeforening 2020. Lastet ned fra tidsskriftet.no