



Hvem er verst?

MEDISIN OG TALL

ARE HUGO PRIPP

E-post: apripp@ous-hf.no

Are Hugo Pripp er forsker og biostatistiker ved Oslo senter for biostatistikk og epidemiologi, Forskningsstøtteavdelingen, Oslo universitetssykehus. Han er professor II ved Fakultet for helsevitenskap, OsloMet – storbyuniversitetet.

Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

Tidlig lærer vi forskjellen på best og verst. Det er derfor ikke overraskende at vi også rangerer medisinsk praksis og helseinstitusjoner slik. Å presentere en rekkefølge er intuitivt og gir innsikt, men den bør tolkes og brukes med varsomhet.

Kvaliteten på helsetjenester opptar oss. Målinger av f.eks. behandling og overlevelse blir sammenlignet og vurdert. Ikke sjelden blir målingene, ofte betegnet kvalitetsindikatorer, presentert fra best til verst (1). Rangeringen kan ha stor betydning for både helseinstitusjoner og enkeltpersoner. Det viktig å ha innsikt i hvordan slike data bør analyseres og presenteres.

Hva gir en forskjell?

I statistikken spiller tilfeldig variasjon ofte en viktig rolle. En rangering uten en analyse av den tilfeldige variasjonen kan gi resultater der vi ikke skiller mellom tilfeldigheter og reelle forskjeller (2). Skal vi rangere sykehus mot hverandre, er to forhold avgjørende: Hvor pålitelig er estimatet for hvert sykehus (usikkerheten innad i sykehuset), og hvor stor er forskjellen mellom sykehusene (heterogeniteten mellom sykehusene)? En kvalitetsindikator er godt egnet til rangering hvis den tilfeldige variasjonen bare forklarer en liten del av de observerte forskjellene mellom sykehusene (3).

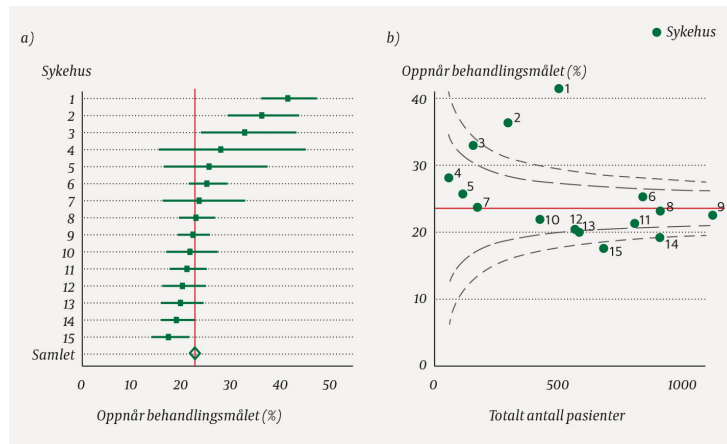
Observerte forskjeller kan ha mange årsaker, ikke bare medisinsk og helsefaglig kvalitet. Foruten tilfeldig variasjon kan det være systematiske forskjeller mellom pasientene, registreringsrutiner, målemetoder og andre observerbare og uobserverbare konfunderende variabler ved sykehusene, noe man må kartlegge og analysere.

Statistisk analyse av forskjeller

Relevante statistiske metoder er illustrert for en nasjonal kvalitetsindikator ved behandling av type 1-diabetes: andelen pasienter for hvert sykehus som oppnår behandlingsmålet med HbA1c mindre enn eller lik 53 mmol/mol ved årskontroll (4).

Et balansediagram («forest plot») viser usikkerheten ved sammenligninger av sykehus mot hverandre og i forhold til gjennomsnittsverdien (5). Tre av sykehusene er statistisk signifikant bedre, og to er dårligere enn gjennomsnittsverdien ved et signifikansnivå på 1 % med tilhørende 99 % konfidensintervaller (figur 1a). De fleste sykehusene er imidlertid

hverken signifikant bedre eller verre i forhold til hverandre.



Figur 1 Forskjeller mellom 15 utvalgte norske sykehus for kvalitetsindikatoren andel pasienter som oppnår behandlingsmålet for diabetes, vist i a) balansediagram med 99 % konfidensintervaller og b) trakttdiagram med kontrollgrenser tilsvarende to (---) eller tre standardavvik (---) fra gjennomsnittsverdien (rød linje).

For sykehus med mange pasienter blir estimatet av en kvalitetsindikator mer nøyaktig enn for sykehus med få pasienter. Vi kan derfor forvente større observerte forskjeller mellom sykehus med få pasienter, enn mellom de med mange. Et trakttdiagram («funnel plot») med kontrollgrenser tilsvarende to og tre standardavvik fra gjennomsnittsverdien er en anbefalt metode for å sammenligne helseinstitusjoner (6). Avvik over to standardavvik er en indikasjon, og avvik over tre standardavvik er et klart tegn på en forskjell utover tilfeldig variasjon. Sykehus 1, 2 og til dels 3 hadde klart høyere andel av pasienter som oppnår behandlingsmålet, mens resultatene for sykehus 14 og 15 tyder på en lavere andel som oppnår behandlingsmålet enn det vi forventer av bare tilfeldig variasjon (figur 1b). De andre sykehusene er stort sett innenfor kontrollgrensene. For sykehus med mange pasienter og tilhørende smale kontrollgrenser kan observerte signifikante forskjeller være mindre klinisk relevante. Ved å bruke trakttdiagram unngår vi en uhensiktsmessig og villedende rangering, samtidig som det gir en klar visuell indikasjon på resultater som avviker.

De siste skal bli de første

De vanskeligste sykdomstilfellene blir gjerne behandlet av de beste og mest avanserte helseinstitusjonene. Tar man ikke hensyn til dette, kan de beste få de verste resultatene siden de utfører de vanskeligste oppgavene. Det er visdom i ordene «... de siste skal bli de første, og de første de siste» (Matt 20,16).

LITTERATUR:

1. HelseDirektoratet. Nasjonale kvalitetsindikatorer (NKI). <https://helsenorge.no/kvalitetsindikatorer> Lest 19.2.2019.
2. Leleu H, Capuano F, Nitenberg G et al. Hospital performance based on treatment delays: comparison of ranking methods. *BMJ Qual Saf* 2014; 23: 73–7. [PubMed][CrossRef]
3. van Dishoeck AM, Lingsma HF, Mackenbach JP et al. Random variation and rankability of hospitals using outcome indicators. *BMJ Qual Saf* 2011; 20: 869–74. [PubMed][CrossRef]
4. HelseDirektoratet. Blodsukkerregulering ved type 1 diabetes hos voksne. https://helsenorge.no/kvalitet-seksjon/Sider/Kvalitetsindikatorer-rapporter.aspx?kiid-KI_HbA1c_type1_voksne Lest 19.2.2019.
5. van Dishoeck AM, Looman CW, van der Wilden-van Lier EC et al. Displaying random variation in comparing hospital performance. *BMJ Qual Saf* 2011; 20: 651–7. [PubMed][CrossRef]
6. Spiegelhalter D. Funnel plots for institutional comparison. *Qual Saf Health Care* 2002; 11: 390–1.

Publisert: 5. juni 2019. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.19.0160
© Tidsskrift for Den norske legeforening 2020. Lastet ned fra tidsskriftet.no