

Kronikk

Kan vi stole på målene for sykehusenes produktivitet?

I Norge publiseres årlig oversikter over utvikling i produktivitet og sykehusenes relative kostnadsnivå gjennom SAMDATA-publikasjonene. Målene vekker stor medieinteresse og benyttes aktivt i vurdering og styring av spesialisthelsetjenesten. Måling av produktivitet i sykehus er imidlertid en krevende oppgave, og det er relevant å diskutere både i hvilken grad de mål som presenteres gir et «sant» bilde av ressursutnyttelse i sektoren og i hvilken grad det faktisk er mulig å utarbeide gode produktivetsmål for helsesektoren. Hensikten med denne artikkelen er å beskrive noen av de utfordringer som man møter ved måling av produktivitet, og på bakgrunn av disse vurdere på hvilke områder produktivetsmål kan være til nytte.

Oppgitte interessekonflikter: Ingen

Jon Magnussen

jon.magnussen@ntnu.no
 Institutt for samfunnsmedisin
 Medisinsk teknisk forskningssenter
 Norges teknisk-naturvitenskapelige universitet
 7489 Trondheim

Helsesektoren kjennetegnes ved knapphet på ressurser, og tilgjengelige ressurser må benyttes på en best mulig måte. Gode produktivetsmål vil i denne sammenhengen være viktige av flere grunner: For det første vil beskrivelser av produktivitet som følger helsetjenestens organisatoriske nivåer inngå

som en viktig del av det løpende styringsgrunnlaget (1, 2). For det andre vil det i utformingen av sektoren være viktig med kunnskap om sammenhengen mellom strukturelle forhold og produktivitet. Sentralt vil være sammenhengen mellom sykehusenes størrelse og produktivitet (3, 4) og sammenhengen mellom sykehusenes aktivitetssammensetning og produktivitet (5, 6). For det tredje er det i den løpende utformingen av helsepolitikken viktig å få informasjon om forholdet mellom helsepolitiske virkemidler og sektorens produktivitet. Har eierskap betydning (7, 8), hva skjer med produktiviteten når man endrer finansieringssystemene (9), hva skjer med produktiviteten når man innfører konkurranse (10) og hva skjer med produktiviteten i sektorer som utsettes for et sterkt ytre politisk press for å øke sin produktivitet (11, 12)?

Gode mål på produktivitet vil kunne utgjøre et viktig grunnlag for styring, planlegging og utforming av helsepolitiske virkemidler. Spørsmålet blir hvor pålitelige slike mål i praksis vil være, og også hvordan den informasjonen de inneholder, best kan formidles.

Produktivitet, effektivitet, kvalitet – hva er hva?

Produktivitet defineres som forholdet mellom produksjon og innsatsfaktorbruk. I mange sammenhenger benyttes begrepet «effektivitet» som forholdet mellom faktisk og best oppnåelige produktivitet (1, 9). I praksis betyr dette at produktivetsmålene normeres, slik at maksimal produktivitet blir 100 og nivåer lavere enn dette kan gis en tolking som prosent måloppnåelse. Dersom ressursbruken måles som totale kostnader, er det vanlig å referere til dette som kostnadseffektivitet. Dersom ressursbruken måles som fysiske innsatsfaktorer (arbeidskraft, kapital etc.), benyttes begrepet teknisk effektivitet. Dette vil kunne støte mot den medisinske forståelse av begrepet effektivitet som er nært koblet til begrepet effekt. Noen vil derfor foretrekke å reservere effektivitetsbegrepet til en diskusjon av forholdet mellom ressursbruk og helse, mens produktivetsbegrepet benyttes i diskusjoner av forholdet mellom ressursbruk og helsetjenester. Det vesentlige her er imidlertid ikke ulik bruk av begreper, men at man i analysen klargjør hva produktivitet (eller effektivitet) måles i forhold til.

Det er viktig å forstå hva de to ulike målene uttrykker. Et produktivetsmål som tar utgangspunkt i produksjonen av helsetjenester, fokuserer kun på «dyktigheten» i gjennomføringen av utført aktivitet, mens et produktivetsmål som tar utgangspunkt i produksjonen av (relativ) helse også trekker inn hvor «dyktig» man er i valg av type pasienter. Det siste målet vil dermed både reflektere hvorvidt man gjør de riktige tingene og hvorvidt man gjør tingene riktig.

Et mulig problem når produktivitet måles i forhold til produksjon av tjenester, er forholdet mellom (lav) produktivitet og (høy) kvalitet. Sammenlikninger av produktivitet forutsetter at det ikke er uobserverbare forskjeller i kvalitet. Dersom kvaliteten er vanskelig observerbar, kan kvalitetsforskjeller feiltolkes som forskjeller i produktivitet. Ideelt sett ville man ønsket å ha egne mål på kvalitet som en del av analysene, men inntil slike foreligger, er det nødvendig å ha spesielt fokus på kvalitetsforskjeller som en mulig forklaring på observerte forskjeller i målt produktivitet.

Utfordringer

Mål på produktivitet er nyttig og nødvendig, men nytten vil avhenge både av hvor presise målene er og av hvordan de blir tolket. Ved empiriske analyser av produktivitet møter man betydelige utfordringer knyttet til kvalitet på aktivitetinformasjon (data), kunnskap og forutsetninger om forholdet mellom ressursinnsats og aktivitet (teknologi) og faktisk beregning av produktivetsmålene (metode).

Data

Sykehusene leverer et sett tjenester som kombineres i behandling av ulike grupper av pasienter. Aktivitet kan dermed måles enten som laboratorieprøver, pleietjenester, operasjonstjenester etc., eller som behandlede pasienter. Fordelen med å fokusere på tjenestetyper er at den enkelte aktivitet er avgrenset og (forholdsvis) lett å definere. Ulempen er at sykehus kan ha høy produktivitet i produksjonen av den enkelte tjeneste, men samtidig et for høyt forbruk av tjenester (f.eks. unødvendige prøver) eller lite hensiktsmessig sammensetning av tjenester. En bedre løsning er derfor å måle produktivitet i forhold til fullstendige pasientforløp. I praksis gjøres dette i Norge gjennom DRG-systemet (13). Det er i dag om lag 500 ulike diagnoserelaterte grupper, og to forhold er

viktig ved bruk av disse i analyser av produktivitet. For det første bør pasientsammensetningen innen hver diagnoserelatert gruppe ikke variere mellom sykehus, for det andre bør de kostnadsvektene som brukes når gruppene vektet sammen, reflektere det relative ressursbehovet på en korrekt måte. Ingen av disse forutsetningene er oppfylt i Norge i dag, og det er heller ingen etablert internasjonal konsensus om hvordan man best skal håndtere en situasjon med flere hundre ulike pasientgrupper i måling av produktivitet. Sykehusene utfører også andre aktiviteter enn pasientbehandling; i særlig grad forskning og undervisning. Disse håndteres i dag gjennom å skille ut den antatte ressursbruken knyttet til forskning og undervisning (1). En bedre måte vil være å konstruere egne mål på disse aktivitetene (14, 15).

Den praktiske utfordringen er dermed å etablere et datasett som er korrekt i den forstand at det fanger opp hele aktiviteten og at registreringen skjer etter de samme prinsipper og med den samme kvalitet over tid og mellom ulike enheter. I så måte er det flere utfordringer i forhold til det data-grunnlaget som i dag benyttes:

- DRG-systemet fanger ikke opp alle variasjoner i pasientsammensetning.
- Systemet baseres på sykehusopphold, ikke pasientforløp.
- Kostnadsvektene som ligger til grunn for aggregering, er usikre.
- Varierende kodepraksis kan føre til at samme type pasient «verdsettes» ulikt ved ulike sykehus.
- Den polikliniske aktiviteten mangler klassifiseringssystem.
- Det mangler gode registre for forsknings- og undervisningsaktivitet.

Teknologi

Mål på produktivitet og effektivitet er interessant fordi de sier noe om forholdet mellom faktisk og mulig ressursutnyttelse. Dette krever kunnskap om hva som er mulig. I økonomisk terminologi innebærer dette kunnskap om teknologien, definert som det teoretiske forholdet mellom mulig aktivitet og størrelsen og sammensetningen av innsatsfaktorene. Sentrale problemstillinger er: I hvilken grad er det proporsjonalitet mellom innsatsfaktorbruk og mulig aktivitet? I hvilken grad kan ulike innsatsfaktorer (for eksempel arbeidskraft og utstyr) erstatte hverandre i pasientbehandlingen? I empiriske analyser vil det teoretisk mulige ikke kunne observeres, og måling skjer dermed relativt til en «beste praksis». Ved etablering av beste praksis må det imidlertid gjøres forutsetninger om teknologien, forutsetninger som vil ha betydning for de produktivetsmål som beregnes.

Dette er særlig komplisert fordi sykehusene behandler en rekke ulike pasientgrupper.

Et viktig spørsmål er om det er rimeligere å behandle flere pasientgrupper i samme sykehus (og i så fall hvilke), eller om kostnadene øker ved økende «bredde» i pasientsammensetningen. Spørsmålet om mulige samdriftsfordeler handler både om struktur (spesialisering eller differensiering) og om mulighetene for å bruke kostnadsvekter til å aggregere de diagnoserelaterte gruppene. Det er i litteraturen ikke etablert noen konsensus om hva som er «riktig» aggregeringsnivå, ei heller om formen på den teknologien som skal estimeres. SAMDATA presenterer to ulike mål; en kostnadsindeks som er basert på en aggregering av alle diagnoserelaterte grupper, men rensket for poliklinisk aktivitet (og kostnader), og et produktivetsmål hvor poliklinisk aktivitet inngår. De bygger imidlertid på følgende forutsetninger:

- Det er ingen samdriftsfordeler eller ulemper.
- Det er muligheter for substitusjon mellom leger og annen arbeidskraft.
- Det er konstant skalaavkastning.

Andre forutsetninger kan lett innarbeides, men vil da gi som resultat andre mål på produktivitet for enkelte sykehus.

Metode

Den tredje utfordringen er knyttet til valg av metode for estimering av produktivitetstallene. Dette skal ikke behandles i detalj her, men det er viktig å være klar over at valg av metodisk tilnærming vil kunne påvirke målene på effektivitet. Dette skyldes at ulike metoder krever ulike former for restriksjoner på den teknologien som skal estimeres, og de vil håndtere mulige utvalgs- og målefeil på forskjellige måter (16).

Drøfting av utfordringene

De utfordringer som her er beskrevet, er ikke spesielle for helsesektoren, selv om det særlig på produksiden er store utfordringer knyttet til operasjonalisering av aktiviteten (17). Samlet innebærer imidlertid dette at vi ved måling av produktivitet må foreta en rekke valg som vil være ad hoc i den forstand at de preges av tilgjengelighet på data, usikkerhet knyttet til faktisk teknologi og det faktum at samme problemstilling kan tilnærmes med flere ulike metoder.

De produktivetsmål som presenteres vil være påvirket av disse valgene. I noen grad kan vanlige statistiske metoder benyttes til å vurdere målenes robusthet (18), men det vil allikevel gjenstå et element av skjønn i de valg som gjøres og dermed de resultater som presenteres. Spørsmålet blir da: Hvilken nytte har vi av denne typen mål i forhold til de målsettingene vi innledningsvis skisserte, og i hvilken grad kan vi forbedre de målene som i dag tolkes som produktivetsmål?

Målenes nytte og deres formidling

Etter 15 år med SAMDATA er det en innarbeidet tradisjon i Norge å vurdere produktivitet og effektivitet på institusjonsnivå. Samtidig er det en utbredt skepsis i litteraturen rundt produktivets- og effektivetsanalyser i helsesektoren knyttet nettopp til bruk av mål på produktivitet på et så lavt nivå (16–19). Flere arbeider viser imidlertid at konklusjoner basert på analyser av forskjeller i produktivitet mellom grupper av sykehus (8, 20) synes å være robuste overfor valg som gjøres både mht. operasjonalisering, teknologi og metode. Det samme gjelder konklusjoner knyttet til effekt av helsepolitiske virkemidler (9). Måling av produktivitet har dermed størst nytte når formålet er å gi grunnlag for strukturbeslutninger og valg av helsepolitiske virkemidler og minst nytte som grunnlag for vurdering av enkeltinstitusjoners (enn si avdelingens) produktivitet.

Hvilke implikasjoner gir dette for presentasjon av kostnadsindekser og effektivitetsmål i SAMDATA? En mulig konklusjon er at disse vil være beheftet med så stor usikkerhet at de ikke bør publiseres. I stedet bør bruken av produktivetsmål avgrenses til vurderinger av (om det er signifikante) forskjeller mellom grupper av sykehus, utvikling over tid og effekt av helsepolitiske virkemidler på sektorens produktivitet. Spørsmålet er likevel om det er nødvendig å være så redd for presentasjon av produktivetsindikatorer. At produktivetsmål på institusjonsnivå vil være følsomme for valg av operasjonalisering, teknologiforutsetninger og metode betyr i seg selv ikke at den informasjon man får fra kostnadsindikatorer på sykehusnivå er verdiløs. Det må imidlertid stilles sterke krav til tolking og bruk (og brukerne) av dem. Informasjon om egen utvikling over tid og informasjon om man ligger høyt eller lavt vil sammen med annen kunnskap kunne danne et underlag for styring i og planlegging av helsetjenesten på institusjonsnivå. Verdien ligger imidlertid ikke i å bruke målene som en bekreftelse verken på udyktighet eller fortreffelighet, men i å benytte dem som ett av mange bidrag i forbedrings- og læringsprosesser. Samme utgangspunkt må også eier ha; snarere enn å bruke kostnadsindekser som entydig grunnlag for budsjettering og styring, må utgangspunktet være at store forskjeller kan skjule ulike løsninger på likeartede problemer og at målene derfor kan danne utgangspunkt for læring.

Konklusjon

Måling av produktivitet er en vanskelig øvelse, og mål på produktivitet vil være beheftet med usikkerhet. Dersom disse benyttes slavisk på et for lavt nivå, kan dette gi grunnlag for feilslutninger. Samtidig er mål på produktivitet omfattet med stor og berettiget interesse fra allmennhet, medier og beslut-

ningstakere. Det synes dermed lite fruktbart å «holde» tilbake de målinger som faktisk utføres på det datagrunnlaget som i dag er til stede. Det bør i det fremtidige arbeidet med dette allikevel fokuseres på de muligheter som faktisk er til stede for forbedring:

For det første bør det arbeides systematisk og grundig med vedlikehold og forbedring av kvaliteten på DRG-systemet og de tilhørende kostnadsvektene. Særlig bør kostnadsvektene, som danner grunnlaget både for produktivitetmåling og finansiering, få en kritisk gjennomgang. For det andre bør datagrunnlaget forbedres både på personell- og kapitalsiden. For det tredje bør det vurderes om presentasjonsformen i større grad bør reflektere at dette er estimater med en tilhørende usikkerhet knyttet til muligheten for utvalgsskjevhet og målefeil. Endelig bør man i enda sterkere grad fokusere på formidlingen av produktivitetsmål. Det å innarbeide mål på usikkerhet som en del av presentasjonen har liten hensikt dersom ikke mottakerne (allmennhet, medier og beslutningstakere) forstår hva produktivitetmålene innebærer og (ikke minst) hva de ikke innebærer.

Manuskriptet ble godkjent 7.9. 2005.

Litteratur

1. SAMDATA somatikk. Sektorrapport 2003. Rapport 1/04. Trondheim: SINTEF Helse, 2004.
2. St.prp. nr. 1 (2004–05). For budsjetterminen 2005.
3. Aletras Vassilis H. A comparison of hospital scale effects in short-run and long-run cost functions. *Health Econ* 1999; 8: 521–30.
4. Kjekshus LE, Hagen TP. Ga sammenlånge av sykehus høyere effektivitet? Erfaringer fra Norge i 1990-årene. *Tidsskrift for Velferdsforskning* 2003; 5: 2–13.
5. Kittelsen SAC, Magnussen J. Economics of scope in Norwegian hospital production – a DEA analysis. HERO Working paper 2003: 8. Oslo: Health Economics Research Programme at the University of Oslo, 2003.
6. Prior D, Solà M. Technical efficiency and economics of diversification in health care. *Health Care Manag Sci* 2000; 3: 299–307.
7. Valdmanis V. Ownership and technical efficiency of hospitals. *Med Care* 1990; 28: 552–61.
8. Mobley LR, Magnussen J. An international comparison of hospital efficiency: does institutional environment matter? *Applied Economics* 1998; 30: 1089–100.
9. Bjørn E, Iversen T, Hagen T et al. The effect of activity-based financing on hospital efficiency: A panel data analysis of DEA efficiency scores 1992–2000. *Health Care Manage Sci* 2003; 6: 271–83.
10. Gift TL, Arnould R, DeBrock L. Is healthy competition healthy? New evidence of the impact of hospital competition. *Inquiry* 2002; 39: 45–55.
11. Halsteinli V, Kittelsen SAC, Magnussen J. Productivity growth in Norwegian psychiatric outpatient clinics for children and youths. *J Ment Health Policy Econ* 2005; 8: 183–191.
12. Halsteinli V, Kittelsen SAC, Magnussen J. Scale, efficiency and organization in Norwegian psychiatric outpatient clinics for children. *J Mental Health Policy Econ* 2001; 4: 79–90.
13. Nordiskt center för klassifikationer i hälso- och sjukvården. www.nordclass.uu.se/index_e.htm
14. Piro F. Påvirker forsknings- og undervisningsinnsatsen kostnadseffektiviteten ved norske sykehus? Analyse av paneldata 1999–2001. Working paper 2004: 4. Oslo: Health Economics Research Programme at the University of Oslo, 2004.
15. Kittelsen SAC, Magnussen J, Piro F. Hva betyr forskning, utdanning og reisetid for sykehusenes kostnader? Working paper 2002: 18. Oslo: Health Economics Research Programme at the University of Oslo, 2002.
16. Newhouse JP. Frontier analysis: how useful a tool for health economics? *J Health Econ* 1994; 13: 317–22.
17. Magnussen J. Efficiency measurement and the operationalization of hospital production. *Health Serv Res* 1996; 31: 21–37.
18. Street A. How much confidence should we place in efficiency estimates? *Health Econ* 2003; 12: 895–907.
19. Folland ST, Hofer RA. How reliable are hospital efficiency estimates? Exploiting the dual to homothetic production. *Health Econ* 2001; 10: 683–98.
20. Valdmanis V. Ownership and technical efficiency of hospitals. *Med Care* 1990; 28: 552–61.