

Online-vurderinger

– sammenheng med systematiske brukerundersøkelser og kliniske målinger?

Notat fra Kunnskapsenteret
Desember 2013

| | |
|-------------------------|---|
| Tittel | Online-vurderinger – sammenheng med systematiske brukerundersøkelser og kliniske målinger? |
| English title | Online ratings – associations with systematic user experience surveys and clinical measures? |
| Institusjon | Nasjonalt kunnskapssenter for helsetjenesten |
| Ansvarlig | Magne Nylenna, direktør |
| Forfattere | Danielsen, Kirsten, forsker, Kunnskapssenteret Kjøllestad, Johanne Gran, forsker, Kunnskapssenteret Bjertnæs, Øyvind Andresen, seksjonsleder, Kunnskapssenteret |
| ISBN | 978-82-8121-784-3 |
| Notat | Desember – 2013 |
| Prosjektnummer | 9971 |
| Publikasjonstype | Notat |
| Antall sider | 21 (24 inklusiv vedlegg) |
| Oppdragsgiver | Intern bestilling |

Sitering Danielsen K, Kjøllestad GK, Bjertnæs, ØA. Online-vurderinger – sammenheng med systematiske brukerundersøkelser og kliniske målinger? Notat 2013. Oslo: Nasjonalt kunnskapssenter for helsetjenesten, 2013.

Kunnskapssenteret takker Legelisten.no for informasjon om Legelisten.no og innspill om relevante nettsteder internasjonalt. Kunnskapssenteret er ansvarlig for innholdet i rapporten.

Nasjonalt kunnskapssenter for helsetjenesten fremskaffer og formidler kunnskap om effekt av metoder, virkemidler og tiltak og om kvalitet innen alle deler av helsetjenesten. Målet er å bidra til gode beslutninger slik at brukerne får best mulig helsetjenester.

Kunnskapssenteret er formelt et forvaltningsorgan under Helse- direktoratet, men har ingen myndighetsfunksjoner og kan ikke instrueres i faglige spørsmål.

Nasjonalt kunnskapssenter for helsetjenesten
Oslo, desember 2013

Oppsummering

Innledning

I løpet av det siste tiåret har det vært en stor økning av nettsider hvor pasienter kan vurdere ulike helsetjenester (online-vurderinger), kvantitativt og/eller kvalitativt. Online-vurderinger har ikke samme forskningsmessige basis som systematiske brukererfaringsundersøkelser, og har blitt kritisert fra ulike hold. Nasjonalt kunnskapssenter for helsetjenesten (Kunnskapssenteret) gjennomførte i 2013 et prosjekt hvor målet var å identifisere og beskrive vitenskapelige studier om sammenhengen mellom online-vurderinger og systematiske kvalitetsmålinger, inkludert systematiske brukererfaringsundersøkelser og kliniske indikatorer. Formålet er å bidra til forskningsbasert kunnskap i debatten om nytten av online-vurderinger, som grunnlag for diskusjoner om, og i så fall hvordan, slike nettsteder kan bidra til økt brukerorientering av norsk helsetjeneste.

Metode

Vi søkte systematisk etter litteratur i følgende databaser:

- OVID Embase 1974 – august 2013
- OVID Medline in-process & other non-indexed 1946 – august 2013
- OVID Medline 1946 – august 2013
- OVID PsycINFO 1987 – august 2013

I tillegg søkte vi i Google og PubMed. Vi søkte også i referanselister til relevante artikler for å utfylle litteratursøket.

Resultater

Søket resulterte i 748 referanser, hvorav 28 ble vurdert for inklusjon og 3 av disse ble vurdert som relevante ut fra inklusjonskriteriene.

Oppsummert viser de tre studiene at det var en sammenheng mellom online-vurderinger og systematiske pasienterfaringsundersøkelser, men sammenhengen var svak til moderat. Sammenhengen mellom online-vurderinger og kliniske mål var

svakere enn sammenhengen mellom online-vurderinger og systematiske brukerundersøkelser.

Konklusjon

Vi konkluderer med at de tre studiene viste at online-vurderinger har svak til moderat korrelasjon med systematiske brukerundersøkelser, og svak korrelasjon med kliniske resultatmål. Etersom det kun ble identifisert tre relevante studier er det behov for mer forskning. Online-vurderinger bør suppleres med resultater fra systematiske brukerundersøkelser og andre kvalitetsmålinger for å gi et riktigst mulig bilde av kvaliteten på helsetjenester på lavt nivå. Det har vært en stor vekst av online-vurderinger internasjonalt, og man må forvente ytterligere ekspansjon både internasjonalt og i Norge i tiden framover.

Innhold

| | |
|---|-----------|
| OPPSUMMERING | 3 |
| Innledning | 3 |
| Metode | 3 |
| Resultater | 3 |
| Konklusjon | 4 |
| INNHold | 5 |
| INNLEDNING | 6 |
| Legelisten.no | 6 |
| Nettsider i utlandet | 7 |
| METODE | 10 |
| Litteratursøking | 10 |
| Inklusjonskriterier | 10 |
| Artikkelutvelging | 11 |
| RESULTAT | 12 |
| Resultat av søk | 12 |
| Artikkel 1, Greaves et al., 2012 | 13 |
| Artikkel 2, Greaves et al., 2012 | 14 |
| Artikkel 3, Bardach et al., 2013 | 15 |
| Annen litteratur om temaet | 16 |
| Enkle empiriske analyser av Legelisten.no | 17 |
| AVSLUTNING | 18 |
| REFERANSELISTE | 20 |
| VEDLEGG 1 | 22 |
| Søkestrategi | 22 |
| VEDLEGG 2 | 24 |
| Spørreskjema fra Legelisten.no | 24 |

Innledning

I løpet av det siste tiåret har det vært en stor økning av nettsider hvor pasienter kan vurdere ulike helsetjenester (online-vurderinger), kvantitativt og/eller kvalitativt. Online-vurderinger har ikke samme forskningsmessige basis som systematiske brukererfaringsundersøkelser, og har blitt kritisert fra ulike hold. Det er blant annet bekymring for om online-vurderinger er representative, og om de gir urimelige negative vurderinger av legene. Samtidig hevder andre at mange muligens undervurderer betydningen av online-vurderinger, da disse kan være med å belyse viktige aspekter ved brukererfaringer med helsetjenester.

Nasjonalt kunnskapssenter for helsetjenesten (Kunnskapssenteret) gjennomførte i 2013 et prosjekt hvor målet var å identifisere og beskrive vitenskapelige studier om sammenhengen mellom online-vurderinger og systematiske kvalitetsmålinger, inkludert systematiske brukererfaringsundersøkelser og kliniske indikatorer. Formålet er å bidra til forskningsbasert kunnskap i debatten om nytten av online-vurderinger, som grunnlag for diskusjoner om, og i så fall hvordan, slike nettsteder kan bidra til økt brukerorientering av norsk helsetjeneste.

I dette notatet vil resultatene fra litteratursøket presenteres. I tillegg gjør vi enkle empiriske analyser av Legelisten.no, for å øke relevansen av notatet for norske forhold. Innledningsvis gis en kort beskrivelse av noen utvalgte nettsider, og vi starter med det norske nettstedet Legelisten.no.

Legelisten.no

Legelisten.no ble lansert i mai 2012 (1). Målet var da å gjøre det enklere å finne en god fastlege. I 2013 ble siden utvidet til også å dekke tannleger. Som pasient kan man gi vurderinger av fastlegen eller leger man tidligere har vært i kontakt med. Dersom man ønsker å gi en vurdering, skriver man en tekstbeskrivelse av den legen man vurderer, og man svarer på ett obligatorisk spørsmål om hvor fornøyd man totalt sett er med legen. I tillegg har man mulighet til å svare på ti valgfrie spørsmål. Disse spørsmålene handler om totalinntrykk, tilgjengelighet, tillit og kommunikasjon, og service. Svarene samles til ulike indekser, og resultatene visualiseres med en skår fra én til fem stjerner. Av bakgrunnsinformasjon svarer

man på spørsmål om alder, kjønn og besøkshyppighet. (Se spørreskjema i vedlegg 2.)

Alle vurderinger blir lest gjennom, og kun de som er i henhold til Legelisten.no sine retningslinjer blir publisert. Retningslinjene inneholder blant annet krav om at krenkende kommentarer ikke blir publisert. I tillegg må vurderingene inneholde en begrunnelse, ikke bare et utsagn som «god» eller «dårlig». Anklagelser om feilbehandling eller feildiagnostisering godtas heller ikke. Kommentarene kan ikke være basert på annet enn egne erfaringer. Dersom man er lege eller tannlege, kan man ikke skrive om seg selv eller andre leger/tannleger ved samme kontor. Familie eller konkurrenter kan heller ikke gi vurderinger. I tillegg finnes en flaggefunksjonalitet hvor alle brukere kan flagge vurderinger de mener ikke hører hjemme på Legelisten.no (2).

Ifølge Legelisten.no var det i november 2013 totalt kommet inn rundt 31 000 vurderinger. Av disse har ca. 23 000 blitt publisert. Det er i gjennomsnitt 5,3 vurderinger per fastlege. Om lag 80 prosent av vurderingene er av positiv karakter (fire eller fem stjerner), og brukerne svarer i snitt på ca. 75 prosent av de frivillige spørsmålene.

Nettsider i utlandet

Internasjonalt finnes et stort omfang av ratingsider hvor man kan gi vurderinger av ulike helsetjenester. I USA drives slike sider av kommersielle aktører (3), og de mest anerkjente ratingsidene er Yelp.com, Angie's List og Rate MDs. I England finnes National Health Service Choices (NHS Choices). I motsetning til i USA drives dette nettstedet av det offentlige.

I USA er online-vurderinger et aktuelt tema i media og flere større tidsskrifter og aviser har diskutert ulike perspektiver ved bruk av slike vurderinger. Blant annet har CBS Evening News stilt spørsmål om man kan stole på slike online-vurderinger (4). Andre perspektiver som ofte blir belyst er legers bekymring for denne typen vurderinger. NBC News skriver i sin artikkel at online-vurderinger ofte enten er veldig positive eller veldig negative. For leger kan negative vurderinger være svært uheldig, og i artikkelen nevnes det at enkelte leger er så bekymret for online-vurderinger at de får pasientene sine til å underskrive kontrakter om at de ikke må legge ut online-vurderinger av dem (5). American Medical News har tatt opp noe av det samme, men skriver at vurderingene som gis ofte er positive fordi det ofte er de fornøyde pasientene som velger å gi en vurdering av legen sin (6). Washington Post har også skrevet om noe av dette i en artikkel som handler om at mange leger frykter å bli vurdert på internett, men at dette er en unødvendig bekymring da det viser seg at vurderingene som gis ofte er positive (7).

Yelp (yelp.com)

Yelp ble grunnlagt i 2004, og målet med nettsiden var å hjelpe folk med å finne gode lokale bedrifter som tannleger, frisører og mekanikere. I dag blir også ulike helsetjenester vurdert på nettsiden. Yelp har vokst raskt i omfang, og hadde i 2013 et gjennomsnitt på ca. 117 millioner månedlige unike besøkende. På nettsiden finnes det over 47 millioner vurderinger. Vurderingene gjøres ved bruk av stjerner på en skala fra én til fem. I tillegg kan man skrive tekstbeskrivelser av det man ønsker å gi vurderinger på.

Angie's List (angieslist.com)

Angie's List er et nettsted som krever medlemskap og innlogging. Det gis vurderinger som handler om 720 ulike tjenester, blant annet helsetjenester. Nettstedet reklamerer for at vurderingene har høy kvalitet og at de er pålitelige. Om lag 60 000 vurderinger blir sendt inn hver måned, og vurderingene gis på en skala fra A til F.

Vurderingene er ikke innlevert anonymt, og informasjon om de som gir vurderingene er tilgjengelig for selskapene som har blitt vurdert. Som begrunnelse for dette, skriver Angie's List at det alltid er to sider av samme sak, og av den grunn får bedriftene/tjenestene mulighet til å svare på vurderingene. Dette skal være med å bidra til at medlemmene får all den informasjonen de trenger før de skal ta en beslutning. Som medlem får man også mulighet til å ta kontakt med Angie's List dersom man ikke får den kvaliteten på tjenesten som er forventet, og Angie's List vil hjelpe til med å løse situasjonen.

Rate MDs (ratemds.com)

Dette er en nettside hvor leger og sykehus blir vurdert. Dersom man ønsker å gi en vurdering, gir man en skår fra én til fem på hva man synes om personalet, ventetid, hjelpsomhet og kunnskap. I tillegg kan man skrive en tekstbeskrivelse av legen/sykehuset. Vurderingene visualiseres med smilefjes (☹ -> ☺), i tillegg til en tallskår fra én til fem på punktene nevnt over. På førstesiden på Rate MDs kommer tekstbeskrivelsene fortløpende opp etter hvert som de skrives inn, og man kan klikke seg direkte inn på den legen eller det sykehuset som blir vurdert.

NHS Choices (nhs.uk)

National Health Service Choices (NHS Choices) er Storbritannias største helsebaserte nettside, og drives av det offentlige. Formålet med nettsiden er å formidle den informasjonen man trenger for å ta helsemessige valg om alt fra livsstilsendringer til valg av offentlige helsetjenester. På NHS Choices kan man både finne online-vurderinger i tillegg til resultater fra tradisjonelle pasienterfaringsundersøkelser. Dersom man ønsker å gi en online-vurdering svarer man på seks spørsmål som omhandler telefontilgjengelighet, om man anbefaler stedet/legen til venner og familie, om man blir møtt med respekt og om

legen/helsetjenesten ivaretar brukermedvirkning. Svarene man gir blir samlet til en skåre som går fra en til fem, og som visualiseres med én til fem stjerner.

Andre aktuelle nettsider

- iwantgreatcare.com
- doctoralia.com
- zocdoc.com
- vitals.com
- healthgrades.com
- jameda.de
- imedo.de
- med.de
- docinsider.de
- artz-auskunft.de

Metode

Vi utførte et innledende litteratursøk i PubMed og Google for å finne søkeord. Tekst- og emneord ble gjennomgått og inkludert i den endelige søkestrategien, som er presentert i vedlegg 1.

I tillegg til litteraturgjennomgangen gjorde vi enkle empiriske analyser av Legelisten.no. Av de i alt 473 fastlegene i Oslo som hadde online-vurderinger på Legelisten.no per 7. november 2013, trakk vi et tilfeldig utvalg på 47 fastleger (10 prosent), og undersøkte antall vurderinger per lege og variasjon i resultater mellom legene. Vi sammenlignet også resultater for fem fastleger fra Legelisten.no med resultater fra en systematisk, men ikke-publisert pasienterfaringsundersøkelse som Kunnskapssenteret har gjennomført.

Litteratursøking

Vi søkte systematisk etter litteratur i følgende databaser:

- OVID Embase 1974 – august 2013
- OVID Medline in-process & other non-indexed 1946 – august 2013
- OVID Medline 1946 – august 2013
- OVID PsycINFO 1987 – august 2013

I tillegg søkte vi i Google og PubMed. Vi søkte også i referanselister til relevante artikler for å utfylle litteratursøket.

Inklusjonskriterier

- Publikasjoner om sammenhengen mellom online-vurderinger og systematiske kvalitetsmålinger, inkludert brukererfaringsundersøkelser og kliniske indikatorer.
- Publikasjoner på engelsk eller et skandinavisk språk.

Artikkelutvelging

To forskere, Johanne Gran Kjøllesdal (JGK) og Kirsten Danielsen (KD), gikk gjennom alle titler og sammendrag for å vurdere relevans i henhold til inklusjonskriteriene. Vurderingene gjorde de uavhengig av hverandre og sammenlignet i etterkant. Der det var uenighet om vurderingene, ble inklusjon eller eksklusjon avgjort ved konsensus. Utvelging av litteratur ble kun gjort basert på tittel og sammendrag.

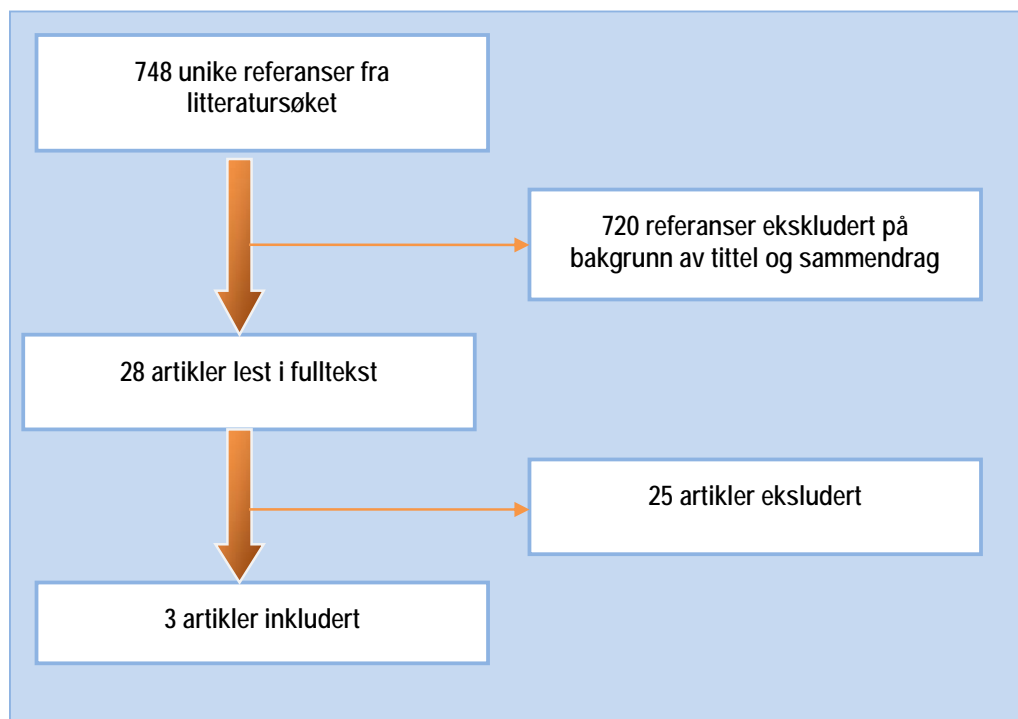
Resultat

Resultat av søk

Søket resulterte i 748 referanser, hvorav 28 ble vurdert for inklusjon og 3 av disse ble vurdert som relevante ut fra inklusjonskriteriene. Disse artiklene sammenlignet online-vurderinger med systematiske kvalitetsmålinger, inkludert brukererfaringsundersøkelser og kliniske indikatorer. I tillegg fant vi noen artikler som vi beholdt som bakgrunns litteratur.

Referanselesing fra fulltekstpublikasjoner ga noen referanse vi vurderte for relevans. Alle referansene ble samlet i en EndNote-database. Etter gjennomgang av samtlige referanser innhentet vi mulig relevante artikler for gjennomlesning av fulltekst.

Figur 1. Flytdiagram med oversikt over litteratursøk og vurdering for inklusjon.



I denne delen av notatet presenterer vi funn fra litteratursøket. Artikler som omhandler vurderinger av sammenhengen mellom online-vurderinger og systematiske brukerundersøkelser og kliniske mål blir beskrevet i korte sammendrag av artiklene. Deretter følger en oppsummering av bakgrunns litteratur vi fant i søket.

Artikkel 1, Greaves et al., 2012

Forfatterne ønsket å undersøke online-vurderinger gitt på nettstedet NHS Choices for 146 sykehus med akuttfunksjon i England. De beskrev svarmønstre i kommentarene, og analyserte sammenhenger med tradisjonelle mål på pasienterfaringer hentet fra en pasienterfaringsundersøkelse. Videre sammenlignet de styrken på sammenhenger mellom online-vurderinger og kliniske utfallsmål, og for pasienterfaringsundersøkelsen og kliniske utfallsmål (3).

Metode

De hentet til dette formålet data fra ulike steder. Pasientenes online-vurderinger ble aggregert til sykehusnivå (trust-nivå) og hentet fra NHS Choices. Det var vurderinger som ble postet i årene 2009/2010. I alt ble 9997 vurderinger tatt med i undersøkelsen. Tradisjonelle pasienterfaringsdata fikk de fra The National Health Services (NHS) undersøkelse blant pasienter som har vært innlagt i somatiske sykehus. Også dette var svar fra pasienter som hadde vært innlagt på sykehus i årene 2009/2010. I alt svarte 68 594 (52 prosent) pasienter på undersøkelsen, som er en årlig pasienterfaringsundersøkelse hvor spørreskjema blir sendt i posten til pasientene.

Tre typer utfallsmål ble brukt i analysene: standardiserte dødelighetsrater, standardiserte dødelighetsrater for høyrisikotilstander, og akutte reinnleggelser innen 28 dager etter utskrivelse fra sykehus. Dataene ble hentet fra Dr. Foster og NHS Comparators.

Det ble gjort deskriptive analyser av de kvantitative dataene fra NHS Choices. Middelerverdier og median ble regnet ut for hver kategori, og mål på spredning inkluderte interkvartil variasjonsbredde. Det totale antallet online-vurderinger for alle sykehusene ble sammenlignet med antall sykehusinnleggelser og antall svar i den nasjonale pasienterfaringsundersøkelsen. Korrelasjonen mellom online-vurderinger og andre mål på pasienterfaringer ble undersøkt ved Spearman's korrelasjonskoeffisient. For å vurdere styrken på sammenheng mellom pasienterfaringer, både online-vurderinger og surveyresultater, og de tre utfallsmålene, ble også Spearman's korrelasjonskoeffisient brukt. Styrken på sammenhengene ble målt ved Fischer's z-transformasjoner.

Resultater

Greaves og kolleger sammenlignet mål på pasienterfaringer fra den spørreskjemabaserte undersøkelsen med pasientvurderingene gitt online. Korrelasjonen mellom online-vurderinger og resultater fra spørreskjemaundersøkelsen (Spearman's korrelasjonskoeffisient) var $\rho = 0,31-0,49$, $p < 0,001$ for alle variabler i analysen. Sammenhengen mellom pasienterfaringer og kliniske utfallsmål viste seg å være minst like sterke for online-vurderinger som for tradisjonelle pasienterfaringsundersøkelser.

Konklusjon

Undersøkelsen oppsummeres med at online-vurderinger kan gi skjevheter i utvalg og resultater, men at de korrelerer med tradisjonelle pasienterfaringsundersøkelser. Online-vurderinger bør ikke erstatte tradisjonelle undersøkelser, men kan gi relevant informasjon om sykehuskvalitet, mer som et supplement til surveydata, og kan være nyttige for pasienter når de skal velge helsetjenester.

Artikkel 2, Greaves et al., 2012

Formålet med studien var å beskrive bruken av NHS Choices, et nettsted der pasienter har mulighet til å vurdere legekantoret de bruker i England. Forfatterne ønsket også å se nærmere på sammenhengen mellom online-vurderinger og tradisjonelle mål på pasienterfaringer og kliniske kvalitetsmål i primærhelsetjenesten (8).

Metode

Online-vurderinger som ble lagt ut på NHS Choices i perioden oktober 2009 til desember 2010, ble aggregert til legekantornivå og brukt i undersøkelsen. Til sammen var det lagt ut 16 952 vurderinger i denne perioden. De tradisjonelle pasienterfaringsdataene online-vurderingene ble sammenlignet med, ble hentet fra undersøkelsen *the General Practice Patient Survey*, en nasjonal spørreskjemabasert undersøkelse der man måler pasienterfaringer på legekantornivå. I de aktuelle årene svarte mer enn 2,1 millioner (37 prosent) pasienter på undersøkelsen.

Med disse dataene som utgangspunkt, undersøkte Greaves og kolleger sammenhenger ved hjelp av regresjonsanalyser og Spearman's korrelasjonskoeffisient. Spørsmål fra den nasjonale spørreundersøkelsen som var likest ratingspørsmålene på NHS Choices, ble brukt. For å kunne sammenligne vurderingene fra nettsiden med mål på klinisk kvalitet, ble andelen pasienter som ville anbefale legekantoret på nettsiden til andre og tradisjonelle kvalitetsmål sammenlignet. Når det gjaldt vurderingene gitt på en skala, ble gjennomsnittet for hvert legekantor regnet ut.

Resultater

I alt var 61 prosent av de 8089 legekantorene i England vurdert, og 69 prosent av vurderingene ville anbefale legekantoret til andre. Legekantor med større andel yngre pasienter, færre dårligstilte pasienter og pasienter bosatt i tettere befolkede områder, hadde større sannsynlighet for å bli vurdert på nettsiden. Det viste seg å være moderate sammenhenger mellom online-vurderinger og pasienterfaringer fra spørreskjemaundersøkelsen (Spearman ρ 0,37 – 0,48, $p < 0,001$ for alle fem variabler), mens det kun ble funnet svake sammenhenger mellom online-vurderinger og kliniske kvalitetsmål (Spearman $\rho \pm 0,18$, $p = < 0,001$ for seks av syv variabler).

Konklusjon

Artikkelforfatterne konkluderer med at hyppigheten av online-vurderinger av allmennleger er varierende i England, men at denne typen vurderinger vanligvis er positive og har en moderat sammenheng med andre mål på pasienterfaringer, samt at vurderingene kun har en svak sammenheng med mål på klinisk kvalitet. Selv om online-vurderinger potensielt kan være mangelfulle og feilaktige, kan de bidra til å lære noe om organisering av tjenestene, og etter hvert som online-vurderinger blir mer og mer vanlig, være en annen måte å se på kvalitet i primærhelsetjenesten på.

Artikkel 3, Bardach et al., 2013

Formålet med studien var å sammenligne sykehusskårer hentet fra en av de mest brukte kommersielle websidene i USA, Yelp.com, med mer systematiske mål på pasienterfaringer og utfallsmål. Artikkelforfatterne ønsket også å undersøke hva som kan gi variasjon på kommersielle ratingsider (9).

Metode

For et nasjonalt utvalg amerikanske sykehus ble skårer på online-vurderinger fra Yelp.com sammenlignet med tradisjonelle mål på sykehuskvalitet. Korrelasjoner ble regnet ut for sykehusenes skårer på Yelp.com og følgende kvalitetsmål: sykehusenes prosent beste vurderinger, det vil si en skåre på ni eller ti på en skala fra én til ti, hvor ti er best på et spørsmål om generell tilfredshet hentet fra pasienterfaringsundersøkelsen *the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS)*.

Også sykehusenes individuelle resultater på denne undersøkelsen, eksempelvis sykepleierkommunikasjon og smertekontroll, samt sykehusenes 30-dagers dødelighetsrater og 30-dagers reinnleggelsesrater, ble tatt med i undersøkelsen. Disse dataene ble sammenlignet med skårer brukerne hadde gitt på Yelp.com. Vurderingene brukerne gir blir på denne nettsiden aggregert og gitt fra én til fem stjerner.

Måten de sammenlignet de to tilbakemeldingsmåtene på var ved å se på korrelasjon mellom Yelp-skårer og de ulike resultatene på pasienterfaringsundersøkelsen og de nevnte kvalitetsindikatorerne. For å sammenligne kjennetegn ved sykehus som enten hadde, eller ikke hadde, online-vurderinger på Yelp.com, ble det brukt t-test og kjikvadrattest. Alle korrelasjoner ble vurdert ved Pearson's korrelasjonskoeffisient. For å undersøke betydningen av varierende antall vurderinger på Yelp.com per sykehus, ble det regnet ut vektete korrelasjonskoeffisienter med vekter omvendt proporsjonalt med antallet online-vurderinger.

Forholdet mellom gjennomsnittlig antall stjerner på Yelp.com og enkeltspørsmål fra spørreskjemaundersøkelsen HCAHPS ble vurdert ved å teste gjennomsnittlig prosent positive svar, det vil si de som krysset av i de to mest positive svarkategoriene på HCAHPS.

Resultater

Av de 3796 sykehusene som hadde resultater fra pasienterfaringsundersøkelsen fra HCAHPS, hadde 962 (25 prosent) skårer på Yelp.com. Blant sykehus med flere enn fem vurderinger på Yelp.com, var korrelasjonen mellom Yelp.com og pasienterfaringsundersøkelsen på 0,49 ($p < 0,001$). Andelen høye vurderinger på hvert av områdene fra pasienterfaringsundersøkelsen økte monotont med økende Yelp-skårer ($p \leq 0,001$ for hvert av områdene). Prosentandelen med de høyeste vurderingene på Yelp.com og på HCAHPS var statistisk signifikant korrelert med lavere dødelighet og hadde færre reinnleggelser for noen utvalgte diagnoser.

Konklusjon

Bardach og kolleger konkluderer med at online-vurderingene som blir gitt på Yelp.com kan vise noe av de samme resultatene som kommer fram ved den tradisjonelle pasienterfaringsundersøkelsen HCAHPS.

Annen litteratur om temaet

Noe av litteraturen vi fant i søket handlet om temaet online-vurderinger. En av artiklene så på online-vurderinger på tyskspråklige nettsider. Innholdet på de ulike sidene ble vurdert og konklusjonen var at det er et vidt spekter av informasjon tilgjengelig på nettsider for online-vurderinger av leger (10). En annen artikkel, en litteraturgjennomgang av publikasjoner om ratingsider, fant blant annet at bruken av nettsider for å vurdere leger fremdeles er ganske begrenset, men økende. Videre fant de at én av seks leger har blitt vurdert online, og at nærmere 90 prosent av vurderingene på denne typen nettsider var positive (11). At online-vurderinger ofte er positive var også blant funnene som ble avdekket i en analyse av kommentarer og vurderinger på den tyske nettsiden Jameda (12). Også andre undersøkelser har vist lignende resultater når det gjelder online-vurderinger (13, 14, 15, 16).

Bevisstheten blant folk om at nettsider for online-vurderinger finnes, var også et tema som ble tatt opp i litteraturen. I en spørreundersøkelse var noen av funnene at 32 prosent av 1505 respondenter hadde hørt om denne typen nettsider, mens elleve prosent selv hadde skrevet en online-vurdering av en lege (17). I en annen studie ble det blant annet undersøkt utviklingen i bruken av nettsider for online-vurderinger, og det kom fram at bruken er økende (18).

Om online-vurderinger gir et godt bilde på pasienterfaringer og pasienttilfredshet, ble også diskutert i en artikkel (19). Det ble gjort en systematisk gjennomgang av både vitenskapelige instrument for å måle pasienterfaringer og pasienttilfredshet, og av nettsider for online-vurderinger. Spørsmål på nettsidene ble sammenlignet med dimensjonene i måleinstrumentene om pasienterfaringer og pasienttilfredshet. Konklusjonen i artikkelen var kort oppsummert at det er nødvendig med mer forskning på dette (19).

Enkle empiriske analyser av Legelisten.no

Av de i alt 473 fastlegene i Oslo som hadde online-vurderinger på Legelisten.no per 7. november 2013, trakk vi et tilfeldig utvalg på 47 fastleger (10 prosent). Antall vurderinger per lege varierte fra én til 26. Mange fastleger fikk svært gode tilbakemeldinger. Rundt halvparten av fastlegene fikk en samlet vurdering på 4,5 eller 5, på en skala fra 1 til 5 hvor 5 er best. Et mindretall (17 prosent) av fastlegene hadde en samleskåre på 3 eller mindre: tre hadde en samleskåre på 2,5, fem en samleskåre på 3. Det var store forskjeller i kvalitative kommentar mellom de fastlegene som skåret høyest og lavest.

Kunnskapssenteret har gjennomført en liten pilotundersøkelse blant fem fastleger, som ledd i utvikling og validering av et innsamlingsopplegg for å måle pasienterfaringer med fastleger. Vi sammenlignet resultatene for disse fem fastlegene (upubliserte data) med data fra Legelisten.no. Vi undersøkte nærmere et spørsmål om samlet vurdering, med skala 1-5 hvor 5 er best (spørsmålene var noe ulikt formulert). Tre av fem fastleger hadde høyere skår på Legelisten.no enn i Kunnskapssenterets måling, mens differansen mellom beste og dårligste lege var langt større på Legelisten.no enn i Kunnskapssenterets måling (1,5 versus 0,4). Rangeringen av fastlegene ut fra de to kildene viste noe variasjon: størst avvik var det for en av fastlegene som fikk tredje beste skår ut fra Legelisten.no, og femte beste skår ifølge Kunnskapssenterets måling.

Avslutning

Kunnskapssenteret gjennomførte høsten 2013 et prosjekt hvor målet var å identifisere og beskrive vitenskapelige studier om sammenhengen mellom online-vurderinger og systematiske kvalitetsmålinger, inkludert systematiske brukererfaringsundersøkelser og kliniske indikatorer. Formålet er å bidra til forskningsbasert kunnskap i debatten om nytten av online-vurderinger, som grunnlag for diskusjoner om, og i så fall hvordan, slike nettsteder kan bidra til økt brukerorientering av norsk helsetjeneste.

Det ble gjennomført et litteratursøk, som resulterte i 748 unike referanser. Tre av disse svarte nøyaktig på inklusjonskriteriene. Av disse var det to av artiklene som brukte data fra sykehus, mens én brukte data om allmennleger.

Oppsummert viser de tre studiene at det var en sammenheng mellom online-vurderinger og systematiske pasienterfaringsundersøkelser, men sammenhengen var svak til moderat (3, 8, 9). Sammenhengen mellom online-vurderinger og kliniske mål var svakere enn sammenhengen mellom online-vurderinger og systematiske brukerundersøkelser. Studiene viste også at majoriteten av vurderingene som gis på nettsider er positive, noe som vi gjenfant i våre enkle analyser av Legelisten.no.

Fordelene med online-vurderinger er at slike vurderinger er oppdaterte, lett tilgjengelige og raskt blir publisert. Et annet viktig moment er at online-vurderinger gir resultater på laveste nivå, fordi man kan hente frem vurderinger på den enkelte lege. Forskning viser at det er variasjon mellom leger, både når det gjelder klinisk kvalitet og brukererfaringer (20). I tillegg er denne måten å samle inn pasienterfaringer på kostnadsbesparende i forhold til tradisjonelle undersøkelser.

Det er imidlertid også mange svakheter med online-vurderinger. Den største er målemetoden som brukes. Metodikken i online-vurderinger ligger langt fra gullstandarden, fordi spørreskjemaene som brukes ikke er validerte og dataene verken er case-mix-justert eller vektet. Man har heller ikke kontroll på hvem som svarer, med påfølgende mulighet for seleksjonsskjevhet. I tillegg kan man påvirkes av andres vurderinger, da disse vil være godt synlige når man skal vurdere selv. I forskningsbaserte målinger er det en forutsetning at individene som svarer er

uavhengige av hverandre. Videre er det ofte få vurderinger per lege. I vår enkle analyse av Legelisten.no varierte antall vurderinger fra én til 26. I en norsk studie viste det seg at man på mer enn 80 prosent av spørsmålene måtte ha mer enn 40 svarere per lege for å få en reliabilitet på legenivå på 0,7, et nivå som vil si at 70 prosent av variasjonen mellom legene er reell og 30 prosent relatert til målefeil (21). Det er med andre ord et stort gap mellom antall svarere som er nødvendig på slike nettstedet for å få tilfredsstillende reliabilitet, og hva som er den faktiske situasjonen.

Vi konkluderer med at de tre studiene viste at online-vurderinger har svak til moderat korrelasjon med systematiske brukerundersøkelser, og svak korrelasjon med kliniske resultatmål. Etersom det kun ble identifisert tre relevante studier er det behov for mer forskning. Nettsider bør suppleres med resultater fra systematiske brukerundersøkelser og andre kvalitetsmålinger for å gi et riktigst mulig bilde av kvaliteten på helsetjenester på lavt nivå, slik det for eksempel gjøres på NHS Choices. Det har vært en stor vekst av nettsider internasjonalt, og man må forvente ytterligere ekspansjon både internasjonalt og i Norge i tiden framover.

Referanseliste

1. www.legelisten.no
2. Legelisten.no. 30.01.2013. Presentasjon for Nasjonalt kunnskapssenter for helsetjenesten.
3. Greaves F, Pape UJ, King D, Darzi A, Majeed A, Wachter RM et al. Associations between Internet-based patient ratings and conventional surveys of patient experience in the English NHS: an observational study. *BMJ Qual Saf* 2012; 21:600-605.
4. CBS-Evening News [Internett]. USA. 2009 [hentet 2013-11-13]. Tilgjengelig fra <http://www.cbsnews.com/stories/2008/12/04/eveningnews/main4649082.shtml>
5. Aleccia J. Docs seek to stifle patients' rants on Web sites. [Internett]. USA: NBC News; 2010. [hentet 2013-11-13]. Tilgjengelig fra http://www.nbcnews.com/id/34794632/ns/health-health_care/#.UoXspvLN6R.
6. Dolan PL. Satisfied patients more likely to leave online review. [Internett]. USA: American Medical News; 2011. [hentet 2013-11-13]. Tilgjengelig fra <http://www.amednews.com/article/20110719/business/307199997/8/>
7. Andrews M Washington Post. Doctors may fear being rated, but they often score high with their patients [Internett]. USA: The Washington Post; 2012 [hentet 2013-11-13]. Tilgjengelig fra http://www.washingtonpost.com/national/health-science/doctors-may-fear-being-rated-but-they-often-score-high-with-their-patients/2012/07/02/gJQAe2bcIW_story.html
8. Greaves F, Pape UJ, Lee H, Smith DM, Darzi A, Majeed A et al . Patients' ratings of family physician practices on the internet: usage and associations with conventional measures of quality in the English National Health Service. *J Med Internet Res* 2012;14:e146.
9. Bardach NS, Asteria-Peñaloza R, Boscardin WJ, Dudley RA. The relationship between commercial website ratings and traditional hospital performance measures in the USA. *BMJ Qual Saf* 2013;22:194-202. [Published Online First 23 November 2012]

10. Emmert M, Sander U, Esslinger S, Maryschok M, Schöffski O. Public reporting in Germany: the content of physician rating websites. *Methods Inf Med* 2012;51:112-120.
11. Emmert M, Sander U, Pisch F. Eight questions about physician-rating websites: A systematic review. *J Med Internet Res* 2013;15:e24.
12. Emmert M, Meier F. An analysis of online evaluations on a physician rating website: evidence from a German public reporting instrument. *J Med Internet Res* 2013;15:e157.
13. Black EW, Thompson LA, Saliba H, Dawson K, Black NM. An analysis of healthcare providers' online ratings. *Inform Prim Care* 2009;17:249-253.
14. Kadry B, Chu LF, Gammas D, Macario A. Analysis of 4999 online physician ratings indicates that most patients give physicians a favorable rating. *J Med Internet Res* 2011;13:e95.
15. Lagu T, Hannon NS, Rothberg MB, Lindenauer PK. Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. *J Gen Intern Med* 2010;25:942-946.
16. Lopez A, Detz A, Ratanawongsa N, Sarkar U. What patients say about their doctors online: A qualitative content analysis. *J Gen Intern Med* 2012;27:685-92.
17. Emmert M, Meier F, Pisch F, Sander U. Physician choice making and characteristics associated with using physician-rating websites: Cross-sectional study. *J Med Internet Res* 2013;15:e187.
18. Gao GG, McCullough JS, Agarwal R, Jha AK. A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. *J Med Internet Res* 2012;14:e38.
19. Reimann S, Strech D. The representation of patient experience and satisfaction in physician rating sites. A criteria-based analysis of English- and German-language sites. *BMC Health Serv Res* 2010;10:332.
20. Birkmeyer JD, Finks JF, O'Reilly A, Oerline M, Carlin AM, Nunn AR et al. Surgical skill and complication rates after bariatric surgery. *N Engl J Med* 2013;369:1434-42.
21. Bjertnaes OA, Lyngstad I, Malterud K, Garratt A. The Norwegian EUROPEP questionnaire for patient evaluation of general practice: data quality, reliability and construct validity. *Fam Pract* 2011;28:342-9.

Vedlegg 1

Søkestrategi

| ID | Searches |
|-----------|---------------------------------|
| #1 | (patient adj4 satisfaction |
| #2 | (patient adj4 experience*) |
| #3 | (user adj4 satisfaction) |
| #4 | (user adj4 experience*) |
| #5 | (consumer adj4 satisfaction) |
| #6 | (consumer adj4 experience*) |
| #7 | patient cent* |
| #8 | patient cent* care |
| #9 | (patient adj4 perception*) |
| #10 | (patient adj4 evaluation*) |
| #11 | (patient adj4 outcome) |
| #12 | patient satisfaction survey* |
| #13 | (patient adj4 survey*) |
| #14 | (patient adj4 data) |
| #15 | (patient adj4 questionnaire*) |
| #16 | (survey adj10 research) |
| #17 | (experience adj10 survey) |
| #18 | (patient adj4 assessment*) |
| #19 | health care survey* |
| #20 | healthcare survey* |
| #21 | health service* survey* |
| #22 | health service* research* |
| #23 | Clinical outcome* |
| #24 | (Quality adj4 measure*) |
| #25 | (Quality adj4 indicator*) |
| #26 | Performance measurement* |
| #27 | Performance measurement system* |
| #28 | Performance measure* |
| #29 | (Performance adj4 indicator*) |

| | |
|-----|---|
| #30 | Provider performance* |
| #31 | (Performance adj4 information*) |
| #32 | (Health adj4 outcome*) |
| #33 | Individual clinical performance* |
| #34 | Clinical performance data |
| #35 | (web adj4 patient rating*) |
| #36 | physician rating web* |
| #37 | physician rating site* |
| #38 | (online doctor adj4 review*) |
| #39 | (commercial adj4 website*) |
| #40 | (consumer adj4 rating*) |
| #41 | (rating adj4 physician*) |
| #42 | (rating adj4 hospital*) |
| #43 | (healthcare adj4 rating*) |
| #44 | (health care adj4 rating*) |
| #45 | (commercial adj4 website*) |
| #46 | (social adj4 media website*) |
| #47 | (internet adj4 rating*) |
| #48 | 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 |
| #49 | 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 |
| #50 | 48 and 49 |

Vedlegg 2

Spørreskjema fra Legelisten.no

1. Totalinntrykk (obligatorisk)
 - a) Din beskrivelse av legen (tekstbeskrivelse)
 - b) Samlet vurdering. Hvor fornøyd er du med denne legen? (skala fra 1 til 5)

2. Tilgjengelighet (valgfritt) (skala fra 1 til 5)
 - a) Hvor lenge må du vanligvis stå i telefonkø ved timebestilling?
 - b) Hvor fort får du vanligvis time?
 - c) På legekantoret, hvor lenge må du vanligvis vente før du kommer inn til timen?

3. Tillit og kommunikasjon (valgfritt) (skala fra 1 til 5)
 - a) Stoler du på at legen gir deg medisinske råd og anbefaling om behandling som er til ditt beste?
 - b) Føler du at legen hjelper deg med å forstå din sykdom?
 - c) Føler du at legen lytter til deg og svarer på dine spørsmål?
 - d) Føler du at legen bruker tilstrekkelig tid på deg?

4. Service (valgfritt) (skala fra 1 til 5)
 - a) Hvor fornøyd er du med åpningstidene til legen/legekantoret?
 - b) Hvor fornøyd er du med øvrig personale på legekantoret (høflighet, vennlighet, effektivitet)?
 - c) Hvor fornøyd er du ellers med legekantoret (rent, hyggelig venteværelse, beliggenhet og annet)?

5. Bakgrunnsinformasjon
 - a) Kjønn
 - b) Alder
 - c) Besøkhypighet

Nasjonalt kunnskapssenter for helsetjenesten
Postboks 7004, St. Olavs plass
N-0130 Oslo
(+47) 23 25 50 00
www.kunnskapssenteret.no
Notat: ISBN 978-82-8121-784-3

desember 2013

 kunnskapssenteret