

Tolker leserne forskningsresultater i tråd med forfatterens intensjoner?

Tove Ringerike¹, Claire Glenton², Gunn Elisabeth Vist³, Gro Jamtvedt⁴ og Magne Nylenna⁵

1) Seksjon for legemidler og helseøkonomi, Nasjonalt kunnskapssenter for helsetjenesten

2) Seksjon for global helse, Nasjonalt kunnskapssenter for helsetjenesten

3) Seksjon for forebyggende, helsefremmende og organisatoriske tiltak, Nasjonalt kunnskapssenter for helsetjenesten

4) Avdeling for kunnskapsoppsummering, Nasjonalt kunnskapssenter for helsetjenesten

5) Nasjonalt kunnskapssenter for helsetjenesten

Korrespondanse: Tove Ringerike, E-post: Tove.Ringerike@kunnskapssenteret.no Telefon: 23 25 50 00

SAMMENDRAG

Bakgrunn: Forskningsresultater kan formidles på ulike måter, både gjennom tall, tekst og kombinasjoner av disse. Utover et effektestimert er det i klinisk forskning også ønskelig å si noe om hvor sikker man er på resultatet, for eksempel ved å oppgi konfidensintervall og gjøre en vurdering av kvaliteten på dokumentasjonen. Vår hensikt med denne undersøkelsen var å få en indikasjon på hvordan våre måter å formidle resultater på, blir oppfattet blant våre lesere.

Materiale og metoder: Vi hentet 17 resultatformuleringer fra Kunnskapssenterets rapporter og utførte en nettbasert spørreundersøkelse. Leserene skulle svare på hva de trodde forfatterne forsøkte å formidle. Åpen invitasjon til å delta i undersøkelsen ble sendt som del av Kunnskapssenterets nyhetsbrev til over 3 000 e-postmottakere.

Resultater: Vi mottok 173 utfylte spørreskjemaer. For samtlige formuleringer var det flertall (mer enn 50 %) for den tolkningen som etter vår vurdering var mest i tråd med forfatterens intensjoner, men andelen varierte fra 54,1 % til 93,4 %.

Konklusjon: Respondentenes tolkninger av Kunnskapssenterets resultatformuleringer samsvarte i hovedsak med intensjonene. Variasjonen i tolkning viste imidlertid at vi må fortsette arbeidet med å finne lettfattelige, men likevel tilstrekkelig presise formuleringer.

Ringerike T, Glenton C, Vist GE, Jamtvedt G, Nylenna M. **Do readers interpret research results in line with the authors' intentions?** *Nor J Epidemiol* 2013; 23 (2): 231-236.

ENGLISH SUMMARY

Background: Research results can be communicated in various ways, for instance through numbers, words, and combinations of these. Within clinical research, in addition to presenting the estimate of effect, it is also desirable to say something about our confidence in this result, for example by presenting a confidence interval and making an assessment of the quality of the evidence. The purpose of this study was to assess how the ways in which we communicate results are perceived by our readers.

Material and methods: We selected 17 results-statements taken from reports published by the Norwegian Knowledge Centre for the Health Services, and conducted an online survey. Participants were asked to respond to what they thought the writers were trying to convey. An open invitation to participate in the survey was sent with the biweekly newsletter for the Norwegian Knowledge Centre for the Health Services, which is delivered to over 3,000 e-mail recipients.

Results: We received 173 completed questionnaires. For all statements, the majority (over 50 %) chose the interpretation that in our assessment was most in line with the authors' intentions, but the proportion varied from 54.1 % to 93.4 %.

Conclusion: Respondents' interpretations of the Knowledge Centre's statements of results corresponded largely with the authors' intentions. The variation in interpretations shows that we must continue to work to find formulations that are easy to understand, yet sufficiently precise.

This is an open access article distributed under the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

BAKGRUNN

Resultater fra kvantitativ forskning kan formidles på ulike måter. Noen foretrekker å få informasjonen i tall og tabeller, noen vil ha informasjonen i tekstform, og noen ønsker seg både tall og tekst (1). Selv i de tilfelle-

ne der det er mulig å tallfeste informasjon, for eksempel innen forskning om effekt av tiltak, vil det finnes ulike måter å presentere resultatene (effektestimertene) på. Resultatene kan gis som tid til noe skjer (for eksempel progresjon av sykdom), prosenter, odds ratio, relativ risiko, antall pasienter som må behandles for å

unngå en hendelse ("number-needed-to-treat"), absolutt risikoreduksjon og gjennomsnittsforskjeller. Lesernes forståelse er ikke alltid knyttet til lesernes preferanser, og de ulike måtene å oppgi effektestimater forstås mer eller mindre godt (2,3). For eksempel, ser det ut til at lesernes forståelse av risiko er mer presis om man presenteres for tall istedenfor ord (4) med absolutte istedenfor relative tall og som hyppighet (natural frequencies) istedenfor prosentandeler (2). Undersøkelser i Norge og utlandet har vist at pasienter og helsepersonell deler mange av de samme utfordringene når det gjelder å forstå forskningsresultater (1,5).

Utover selve effektestimater er det i klinisk forskning også ønskelig å si noe om hvor sikker man er på resultatet, for eksempel ved å oppgi 95 % konfidensintervall rundt effektestimater. Svakheter ved måten studier ble gjennomført på er en annen kilde til usikkerhet om effektestimaterets gyldighet. For systematiske oversikter der resultater fra flere studier samles, vil vår tillit til resultatene også påvirkes av antall studier, antall hendelser og størrelsen på studiene, samt hvorvidt alle resultater peker i samme retning. Slike forhold gjør at forskere ikke bare oppgir et effektestimater, men også forsøker å beskrive usikkerheten knyttet til resultatene og hvilken grad av tillit man har til det beregnede effektestimater (kvalitet på dokumentasjonen) (6).

Nasjonalt kunnskapssenter for helsetjenesten (Kunnskapssenteret) har som en av sine hovedoppgaver å sammenstille forskning om effekter av helsetiltak. De tiltak vi utarbeider rapporter på (systematiske oversikter og metodevurderinger), bestemmes i hovedsak etter en årlig bestillingsrunde som er åpen for alle som ønsker svar på viktige problemstillinger innen helsetjenesten. Rapportene brukes som underlagsdokumenter for overordnede helsepolitiske beslutninger, i retningslinjer, prioriteringer og for klinisk praksis. Det er derfor viktig å unngå feil, feiltolkninger og misforståelser. Våre kunnskapssoppsummeringer lages med systematiske metoder, som beskrevet i vår metodebok (7). Rapportene er fritt tilgjengelige og nedlastbare for alle.

Kunnskapssenteret ønsker å presentere forskningsdokumentasjon om effekt av tiltak på en forståelig, nøytral og presis måte. Dette inkluderer også å gi sluttbrukeren informasjon om hvor sikre vi er på den foreliggende dokumentasjonen. Vi bruker graderingsverktøyet GRADE (Grading of Recommendations Assessment, Development and Evaluation) til å systematisere våre vurderinger vedrørende usikkerhet til resultatene (8). Der det finnes usikkerhet, presenteres dette på forskjellige måter blant Kunnskapssenterets rapportforfattere, for eksempel gjennom henvisninger til GRADE-vurderingen ("kvaliteten på dokumentasjonen var lav") eller gjennom modifierende ord ("Behandlingen fører *muligens* til bedring"). Vi vet lite om hvordan slike utsagn oppfattes blant våre lesere.

Det har tidvis kommet tilbakemeldinger på at enkelte av våre resultatformuleringer kan være vanskelige å forstå. Vi utførte derfor en enkel pilotundersøkelse for å undersøke om resultatene Kunnskapssenteret presenterer blir tolket og forstått slik vi ønsker.

METODE

Vi utførte en spørreundersøkelse for å få innblikk i hvordan lesere av Kunnskapssenterets rapporter tolker forskningsresultatene, og hvorvidt tolkningene er forskjellige fra våre egne. Vi utarbeidet et eget spørreskjema som ble diskutert i en utvidet arbeidsgruppe og pilottestet internt i Kunnskapssenteret. Det besto av en introduksjon, spørsmål om arbeidssted/situasjon, resultatformuleringene hentet fra Kunnskapssenterets rapporter, tolkingsalternativer for formuleringene og mulighet for å gi kommentarer.

Utforming av spørreskjema

For å identifisere aktuelle resultatformuleringer, tok vi utgangspunkt i systematiske oversikter og metodevurderinger ("Health Technology Assessments") publisert av Kunnskapssenteret i perioden 2008 til 2011. Kunnskapssenterets oversikter innledes av en side der hovedfunnene presenteres (ofte i kulepunkt), et tre siders utvidet sammendrag, og til slutt den fullstendige kunnskapssoppsummeringen. For denne spørreundersøkelsen hentet vi først ut resultatformuleringene slik de var presentert i hovedfunnene og sammendraget. Så gjorde vi et strategisk utvalg der vi valgte formuleringer som brukte forskjellige måter å presentere resultater på og som representerte forskjellige fagområder. For å få frem ulike mulige tolkninger av formuleringene, valgte vi seks enkle svaralternativer.

Alternativene er vist i figur 1. På bakgrunn av de fullstendige publikasjonene, og dermed med mulighet for utfyllende resultater og omtale, gjorde vi (forfatterne) en vurdering av hvilken hovedkategori vi mente lå nærmest hver resultatformulering, og som vi dermed regnet som korrekt svaralternativ. Oppsett og svaralternativer var faste for alle spørsmålene for å sikre gjenkjenning gjennom undersøkelsen. Oppsettet besto av formuleringen fra Kunnskapssenterets rapport etterfulgt av setningen "Etter å ha lest dette utsagnet tror jeg forskerne bak rapporten har forsøkt å si følgende om tiltaket" og svaralternativene (se figur 1).

Vi ba respondentene oppgi hvilken arbeidsrelasjon de har til Kunnskapssenteret for å kunne undersøke om ulike målgrupper har ulike fortolkninger av våre formuleringer.

Spørreundersøkelsen var anonym og frivillig. Det ble ikke søkt om etiske eller personvernmessige tilatelser til gjennomføring av undersøkelsen.

Nettbasert spørreundersøkelse

Spørreundersøkelsen var nettbasert og ble laget i Google docs (docs.google.com). Resultatene ble eksportert til Excel for videre håndtering. Vi presenterer resultatene som rene optellinger. Fordi et flertall av svarene falt i én kategori valgte vi å ta utgangspunkt i disse (de som svarte det vi mente var mest riktig) og utførte en khi-kvadrat test for å se om det var forskjell mellom de forskjellige respondentgruppene i hvor stor andel som svarte riktig alternativ. Vanligvis settes signifikansnivå til 5 %, men vi har justert for at vi har

Spørreundersøkelse

Spørsmål 11

"Resultatene tyder på at acetylsalisylsyre muligens forebygger hjerteinfarkt, men kvaliteten på dokumentasjonen var lav"

Etter å ha lest dette utsagnet tror jeg forskere bak rapporten forsøker å si følgende om tiltaket

- Dette virker helt sikkert
- Vi tror det virker, men er ikke sikre
- Vi tror det virker like bra
- Vi tror det ikke virker, men er usikre
- Dette virker helt sikkert ikke
- Vi vet ikke

Kommentar til spørsmål 11

[Tilbake](#) [Fortsett](#)

Drevet av [Google Dokumenter](#)

[Rapporter misbruk](#) - [Vilkår for bruk](#) - [Ytterligere vilkår](#)

Figur 1. Bilde av hvordan spørreskjema så ut på nettløsningen.

korrelerte utfall og brukte derfor et nominelt signifikansnivå på 0,29 % (5 % delt på 17 spørsmål, Bonferroni-korreksjon). Informasjon og lenke til spørreundersøkelsen ble sendt ut med Kunnskapssenterets nyhetsbrev den 16. februar 2012. Nyhetsbrevet sendes til ca. 3 000 mottakere. I tillegg gikk en egen e-post til alle ansatte i Kunnskapssenteret og en til ledere i Helsedirektoratet. Vi avsluttet spørreundersøkelsen 27. februar 2012. Det ble ikke sendt påminnelser eller purringer.

RESULTATER

Vi mottok totalt 173 utfylte spørreskjemaer. Egenrapportert tilhørighet fordelte seg med følgende antall respondenter: 50 ansatt i Kunnskapssenteret, 26 ansatt i Helsedirektoratet, 48 var helsepersonell, 11 krysset av for leder/beslutningstaker i helsetjenesten, to definerte seg som pasient og 36 som annet.

Det var noen respondenter som ikke ga svar på alle formuleringene, derfor varierer antall svar per formulering fra 165 til 173. En fullstendig oversikt over de utsagnene som ble brukt i spørreundersøkelsen og hvordan svarene fordelte seg er vist i tabell 1. For samtlige formuleringer var det rent flertall (mer enn 50 %) for den tolkningen vi anså som mest korrekt (uthevet skrift i tabell 1), men andelen varierte fra 54,1 % til 93,4 %.

I tabell 2 presenterer vi andelen korrekte svar for hver av respondentgruppene. Vi påviste ingen statistisk

signifikante forskjeller, bortsett fra ett tilfelle (utsagn 4).

Flere av kommentarene i spørreundersøkelsen tydet på at enkelte av utsagnene var i vanskelig tilgjengelig språk og/eller tungvint formulert. Kommentarer til samme utsagn kunne variere fra at leseren var fornøyd, til bemerkninger om dårlig språk eller at utsagnet var vanskelig å forstå.

DISKUSJON

Våre funn viste at respondentene i hovedsak hadde god forståelse av Kunnskapssenterets resultatformuleringer. Det var godt samsvar mellom det svaralternativet med høyest andel svar og det svaralternativet vi hadde vurdert som den mest korrekte tolkningen, for hver formulering (tabell 1).

Bruk av Kunnskapssenterets nyhetsbrev som kanal for å rekruttere deltakere til undersøkelsen innebærer antakelig en seleksjonsskjevhet da abonnenter på nyhetsbrevet neppe er representative for alle som leser våre rapporter. Den lave svarprosenten (5 % av abonnentene) tilsier både at våre resultater er usikre og at vi ikke uten videre kan vite om resultatene er generaliserbare. Likevel: Undersøkelsen har gitt oss innspill fra over 150 personer, som er en betydelig forbedring av kunnskapsgrunnlaget sammenliknet med de få anekdotene vi hadde fra før av.

Bruk av påminnelser kunne ha gitt oss noen flere

Tabell 1. Oversikt over formuleringer som ble undersøkt og svarfordeling.

Formuleringene er hentet ordrett fra rapporter fra Kunnskapscenteret, med unntak av at vi har lagt inn ordet ”avføringsmiddel” for å forklare hva metylnaltrekson er, spørsmål 6.

Utsagn	Formulering/ Svaralternativ	Dette virker helt sikkert	Vi tror det virker, men er ikke sikre	Vi tror de er like bra	Dette virker helt sik- kert ikke	Vi tror det ikke vir- ker, men er usikre	Vi vet ikke
1	”Vi kan ikke konkludere sikkert, men tilrettelagt dagaktivitetstilbud kan bidra til noe lavere forekomst av atferdsproblemer (fem studier – svært lav kvalitet).”	4 (2 %)	123 (71 %)	4 (2 %)	3 (2 %)	18 (10 %)	21 (12 %)
2	”Det ser ikke ut til at tilrettelagt dagaktivitetstilbud påvirker funksjonsnivået til personer med demens.”	2 (1 %)	5 (3 %)	3 (2 %)	46 (27 %)	108 (63 %)	8 (5 %)
3	”Vitamin B6 og omega-3-fettsyren DHA har muligens ingen effekt på schizofrenisymptomer (lav kvalitet på dokumentasjonen).”	0 (0 %)	0 (0 %)	2 (1 %)	16 (9 %)	132 (77 %)	22 (13 %)
4	”Vi er usikre på effekten av omega-6-fettsyren GLA og vitamin E på schizofrenisymptomer (svært lav kvalitet på dokumentasjonen).”	0 (0 %)	6 (4 %)	6 (4 %)	11 (6 %)	56 (33 %)	93 (54 %)
5	”Behandling av pasienter med akutt forverring av KOLS i ”hjemmesykehus” tenderte mot en reduksjon i risiko for død sammenlignet med behandling i sykehus, men forskjellen nådde ikke statistisk signifikans.”	1 (1 %)	111 (65 %)	19 (11 %)	6 (4 %)	9 (5 %)	25 (15 %)
6	”Resultater fra to randomiserte kontrollerte studier viste at det var flere pasienter med avføring innen fire timer hos de som fikk avføringsmiddelet metylnaltrekson sammenlignet med de som fikk placebo.”	128 (76 %)	32 (19 %)	1 (1 %)	0 (0 %)	2 (1 %)	6 (4 %)
7	”Kognitiv atferdsterapi gir muligens noe økning i arbeidslivsdeltakelse sammenlignet med standard behandling.”	3 (2 %)	142 (83 %)	7 (4 %)	0 (0 %)	12 (7 %)	8 (5 %)
8	”Kvaliteten på dokumentasjonen er for mangelfull til at vi kan trekke konklusjoner om effekt av kosttilskudd og alternativ behandling.”	0 (0 %)	3 (2 %)	1 (1 %)	5 (3 %)	16 (9 %)	145 (85 %)
9	”Det er ikke dokumentert noen forskjeller i effekt og sikkerhet mellom kikkhullskirurgi og vaginal prosedyre ved fjerning av livmor.”	7 (4 %)	3 (2 %)	122 (72 %)	7 (4 %)	5 (3 %)	26 (15 %)
10	”Den systematiske oversikten viste at langtidsbehandling i et terapeutisk samfunn (TS) ikke ga signifikant bedre resultater enn annen type behandling for rusavhengige i døgninstitusjoner.”	7 (4 %)	3 (2 %)	110 (65 %)	26 (15 %)	8 (5 %)	15 (9 %)
11	”Resultatene tyder på at acetylsalisylsyre muligens forebygger hjerteinfarkt, men kvaliteten på dokumentasjonen var lav.”	3 (2 %)	134 (79 %)	1 (1 %)	1 (1 %)	5 (3 %)	25 (15 %)
12	”Vi fant dokumentasjon av høy kvalitet for at behandling med blodtrykksenkende og lipidsenkende legemidler reduserer risikoen for hjerte- og karsykdom.”	152 (91 %)	10 (6 %)	2 (1 %)	1 (1 %)	0 (0 %)	2 (1 %)
13	”Basert på resultater fra én studie ser det ut til at metformin muligens har en effekt på total dødelighet (relativ risikoreduksjon 32 %, 95 % KI 7 % til 51 %).”	5 (3 %)	126 (76 %)	2 (1 %)	1 (1 %)	4 (1 %)	28 (17 %)
14	”Det er liten eller ingen forskjell i fatigue hos brystkreftpasienter som trener under strålebehandling sammenlignet med dem som ikke trener.”	10 (6 %)	4 (2 %)	96 (58 %)	31 (19 %)	15 (9 %)	10 (6 %)
15	”Det er motstridende resultater fra to nyere studier hvorvidt gangtrening reduserer fatigue for pasienter med prostatakreft under strålebehandling.”	0 (0 %)	3 (2 %)	6 (4 %)	2 (1 %)	4 (2 %)	153 (91 %)
16	”Det er solid grunnlag for å si at behandling og rehabilitering i egne slagenheter gir lavere dødelighet sammenliknet med andre måter å organisere slagrehabilitering på.”	156 (93 %)	9 (5 %)	0 (0 %)	0 (0 %)	1 (1 %)	1 (1 %)
17	”Det synes ikke å være forskjeller av betydning mellom ultralydundersøkelse i første eller andre trimester på terminfastsettelse.”	8 (5 %)	7 (4 %)	121 (73 %)	6 (4 %)	12 (7 %)	11 (7 %)

svar, men ville ikke løst seleksjonsskjevheten. Det finnes alternative måter vi kunne valgt ut personer på, for eksempel ved å henvende oss til dem som har bestilt kunnskapsoppsummeringer fra Kunnskapscenteret, eller utvalgte representanter for våre målgrupper.

En relativt stor andel av svarene kom fra ansatte i Kunnskapscenteret. Vi valgte å ikke ekskludere dem ettersom langt fra alle i Kunnskapscenteret lager den type rapporter som vår undersøkelse dreide seg om. Og selv blant dem som gjør det varierer graden av fortrolighet med de ulike måtene å formidle resultater på.

Vi fant ingen statistisk signifikante forskjeller mellom de ulike gruppene av respondenter. Dette var i og for seg ikke overraskende da gruppene vi sammenliknet var så små at statistiske tester har begrenset verdi. Minst halvparten i hver gruppe valgte, med få unntak, korrekt svaralternativ. Det kan tyde på at forståelsen av formuleringene stort sett var rimelig god, i alle gruppene.

I teorien kunne én person svart på spørreundersøkelsen mer enn en gang – vi hadde ikke mulighet til å kontrollere dette. Vi vet heller ikke hvor mange som eventuelt har gått lei underveis og ikke sendt inn skjemmet.

Det er uvisst om vi ville fått samme fordeling mellom svaralternativene dersom vi hadde sammenliknet forskjellige formuleringer for samme resultat. For eksempel inneholder utsagnet ”Resultatene tyder på at acetylsalisylsyre muligens forebygger hjerteinfarkt, men kvaliteten på dokumentasjonen var lav” tre elementer som formidler usikkerhet knyttet til resultatene: ”tyder på”, ”muligens” og ”kvaliteten på dokumentasjonen var lav”. Spørreundersøkelsen var ikke laget for å undersøke tolkninger av ulike formuleringer av samme resultat. En annen undersøkelse kunne ha belyst hvor mange modererende elementer vi må sette inn for å formidle usikkerhet på en tilstrekkelig god måte.

Vi benyttet samme oppsett og svaralternativ til alle

Tabell 2. Andel korrekte svar, for hver respondentgruppe.

Utsagn	Formulering/ Svaralternativ	Totalt (N=173)	Kunn- skaps- senteret (N=50)	Helse- direkto- ratet (N=26)	Helse- personell (N=48)	Leder/ beslut- ningstaker (N=11)	Pasient (N=2)	Annet (N=36)
1	”Vi kan ikke konkludere sikkert, men tilrettelagt dagaktivitetstilbud kan bidra til noe lavere forekomst av atferdsproblemer (fem studier – svært lav kvalitet).”	123 (71 %)	37 (74 %)	20 (77 %)	36 (75 %)	6 (55 %)	1 (50 %)	23 (64 %)
2	”Det ser ikke ut til at tilrettelagt dagaktivitetstilbud påvirker funksjonsnivået til personer med demens.”	108 (63 %)	35 (70 %)	15 (58 %)	25 (53 %)	6 (55 %)	0 (0 %)	27 (75 %)
3	”Vitamin B6 og omega-3-fettsyren DHA har muligens ingen effekt på schizofrenisymptomer (lav kvalitet på dokumentasjonen).”	132 (77 %)	36 (72 %)	23 (89 %)	38 (79 %)	8 (80 %)	1 (50 %)	26 (72 %)
4	”Vi er usikre på effekten av omega-6-fettsyren GLA og vitamin E på schizofrenisymptomer (svært lav kvalitet på dokumentasjonen).”	93 (54 %)	36 (72 %)	13 (50 %)	16 (33 %)	6 (55 %)	0 (0 %)	22 (63 %)
5	”Behandling av pasienter med akutt forverring av KOLS i ”hjemme- sykehus” tenderte mot en reduksjon i risiko for død sammenlignet med behandling i sykehus, men forskjellen nådde ikke statistisk signifikans.”	111 (65 %)	32 (64 %)	16 (62 %)	32 (67 %)	7 (64 %)	2 (100 %)	22 (65 %)
6	”Resultater fra to randomiserte kontrollerte studier viste at det var flere pasienter med avføring innen fire timer hos de som fikk avførings- middelet metylinaltrekson sammenlignet med de som fikk placebo.”	128 (76 %)	38 (78 %)	18 (72 %)	34 (71 %)	5 (50 %)	1 (50 %)	32 (91 %)
7	”Kognitiv atferdsterapi gir muligens noe økning i arbeidslivsdeltakelse sammenlignet med standard behandling.”	142 (83 %)	44 (88 %)	22 (85 %)	33 (69 %)	10 (91 %)	2 (100 %)	31 (89 %)
8	”Kvaliteten på dokumentasjonen er for mangelfull til at vi kan trekke konklusjoner om effekt av kosttilskudd og alternativ behandling.”	145 (85 %)	45 (92 %)	23 (89 %)	38 (79 %)	9 (82 %)	1 (50 %)	29 (85 %)
9	”Det er ikke dokumentert noen forskjeller i effekt og sikkerhet mellom kikkhullskirurgi og vaginal prosedyre ved fjerning av livmor.”	122 (72 %)	39 (78 %)	17 (68 %)	34 (72 %)	6 (55 %)	1 (50 %)	25 (71 %)
10	”Den systematiske oversikten viste at langtidsbehandling i et terapeutisk samfunn (TS) ikke ga signifikant bedre resultater enn annen type behandling for rusavhengige i døgninstitusjoner.”	110 (65 %)	38 (78 %)	16 (62 %)	30 (64 %)	2 (18 %)	1 (50 %)	23 (68 %)
11	”Resultatene tyder på at acetylsalisylsyre muligens forebygger hjerteinfarkt, men kvaliteten på dokumentasjonen var lav.”	134 (79 %)	42 (88 %)	20 (77 %)	36 (77 %)	7 (64 %)	2 (100 %)	27 (77 %)
12	”Vi fant dokumentasjon av høy kvalitet for at behandling med blodtrykksenkende og lipidsenkende legemidler reduserer risikoen for hjerte- og karsykdom.”	152 (91 %)	43 (90 %)	25 (100 %)	43 (92 %)	10 (91 %)	1 (50 %)	30 (88 %)
13	”Basert på resultater fra én studie ser det ut til at metformin muligens har en effekt på total dødelighet (relativ risikoreduksjon 32 %, 95 % KI 7 % til 51 %).”	126 (76 %)	39 (83 %)	15 (60 %)	35 (75 %)	7 (63 %)	2 (100 %)	28 (82 %)
14	”Det er liten eller ingen forskjell i fatigue hos brystkreftpasienter som trener under strålebehandling sammenlignet med dem som ikke trener.”	96 (58 %)	29 (62 %)	16 (62 %)	24 (51 %)	5 (50 %)	1 (50 %)	21 (62 %)
15	”Det er motstridende resultater fra to nyere studier hvorvidt gangtrening reducerer fatigue for pasienter med prostatakreft under strålebehandling.”	153 (91 %)	44 (92 %)	25 (96 %)	42 (89 %)	10 (91 %)	0 (0 %)	32 (94 %)
16	”Det er solid grunnlag for å si at behandling og rehabilitering i egne slagenheter gir lavere dødelighet sammenliknet med andre måter å organisere slagrehabilitering på.”	156 (93 %)	44 (94 %)	25 (96 %)	43 (92 %)	11 (100 %)	1 (50 %)	32 (94 %)
17	”Det synes ikke å være forskjeller av betydning mellom ultralyd- undersøkelse i første eller andre trimester på terminfastsettelse.”	121 (73 %)	38 (81 %)	16 (67 %)	33 (70 %)	5 (46 %)	1 (50 %)	28 (82 %)

utsagnene. Dette kan øke gjenkjenning og lærings-effekt gjennom undersøkelsen. Samtidig medfører det at vi ikke fikk spesialtilpasset svaralternativene til de ulike utsagnene, noe som kan ha forvirret respondene. Dette kan være én av forklaringene på svarfordelingen til utsagnet om trening hos brystkreftpasienter som får strålebehandling: ”Det er liten eller ingen forskjell i fatigue hos brystkreftpasienter som trener under strålebehandling sammenlignet med dem som ikke trener”. Dersom man tar utgangspunkt i at man lurer på om det er minst fatigue med eller uten trening, kan det være nærliggende å tenke at det ikke er vist forskjell mellom disse alternativene, og dermed velger svaralternativet ”Vi tror det virker like bra”. Dersom man i stedet tenker at hensikten er å se om trening er et effektivt tiltak, vil det være nærliggende å velge ”Dette virker helt sikkert ikke”.

Et overraskende funn for oss var at et utsagn vi regnet som nøytralt, ble tolket negativt: Til utsagnet ”Vi

er usikre på effekten av omega-6-fettsyren GLA og vitamin E på schizofrenisymptomer (svært lav kvalitet på dokumentasjonen)” svarte hele 32,6 % at de trodde forskerne forsøkte å si ”Vi tror det ikke virker, men er usikre. Omkring halvparten svarte det vi regnet som mest korrekt: ”Vi vet ikke”. Kanskje er det slik at usikkerhet tolkes mer i retning av at et tiltak ikke har noen effekt, enn motsatt? Andre mulige forklaringer på akkurat denne svarfordelingen kan være at våre lesere har en forutinntatt forestilling om at kosttilskudd/vitaminutskudd ikke har effekt på sykdom, eller at vår rekkefølge på utsagnene har hatt betydning – utsagnet like før dette dreiet seg om to andre kosttilskudd/vitaminutskudd med muligens ingen effekt. Utsagnet kan også brukes for å diskutere forskjellen på ”evidens for ingen effekt” (det virker ikke) og ”ingen evidens for effekt” (vi er usikre på om det virker).

Vi hadde for få og varierende utsagn til å kunne si noe sikkert om mønster eller trender. Det kan likevel

se ut som de som svarte ”feil”, hellet mot en mer negativ tolking enn det vi regnet som ”rett” svar, enten i form av mindre tiltro til resultatet (de trodde forfatterne var mer usikre enn de var) eller mindre tiltro til effekt av tiltakene (like gode i stedet for å tro på en forskjell). Det ser også ut som våre modererende ord gjorde at noen ble mer usikre og la seg på fortolkningen at forfatterne ikke visste (”vi vet ikke”). Muligheten for at modererende ord gjør leserne usikre på om vi faktisk kan si noe om tiltakets effekt eller at det tolkes i mer negativ retning, er noe vi må ha in mente i vårt videre arbeid.

En tidligere undersøkelse utført på systematiske oversikter publisert i Cochrane-biblioteket bekreftet at tolkning av forskningsresultater er utfordrende (9). I motsetning til oss, som undersøkte 17 formidlingssetninger med 173 respondenter, benyttet Ezzo og medarbeidere 160 formidlingssetninger med kun to respondenter til hver. De rapporterte at det var særlig utfordrende å tolke de ikke-signifikante resultatene, og fant i tillegg at de som leste resultatene hadde en tendens til å være mer kritiske i sin tolkning av resultatene enn oversiktsforfatterne selv. Dette er i samsvar med våre funn.

Tross metodesvakheter i vår spørreundersøkelse har vi fått ett innblikk i om våre formuleringer er forståelige for leserne. Det er gledelig at det var såpass godt samsvar mellom respondentenes og vår tolkning av de ulike utsagnene.

Et viktig underliggende spørsmål for Kunnskapssenteret er om enklere formuleringer som dem vi benyttet som tolkningsalternativene i denne undersøkelsen,

bør erstatte de mer innflokke resultatformuleringene som gjerne brukes i våre rapporter i dag. Dersom forenklingene var fullt ut dekkende for vårt budskap, vil det være en naturlig konsekvens. Som regel ønsker vi imidlertid å formidle et mer presist budskap, ofte knyttet opp mot to dimensjoner: 1) hvor stor en eventuell effekt av et tiltak er, og 2) hvor sikre vi er på resultatet. Målet er å gi tilstrekkelig med detaljer til å vise hva forskningsdokumentasjonen egentlig viser av resultater og forbehold, samtidig som vi ønsker at konklusjonen skal være så lettfattelig som mulig. Å avveie disse hensynene er en utfordring for alle som bedriver forskningsformidling.

KONKLUSJON

Tross metodemessige svakheter tyder resultatene fra denne spørreundersøkelsen på at lesernes forståelse av resultatene i våre rapporter i stor grad er i tråd med det vi ønsker å formidle. I flertallet av tilfellene ble utsagnene fra Kunnskapssenterets rapporter tolket i tråd med det vi mente var korrekt. Men det er viktig å merke seg at mange tolket resultatene annerledes enn oss. Vi har altså en viktig utfordring med tanke på å kunne presentere våre resultater slik at leserens tolkning og forståelse er i samsvar med det vi forsøker å formidle. Vi fortsetter derfor vårt arbeid med å finne frem til lettfattelige, men likevel tilstrekkelig presise formuleringer, og vil vurdere å standardisere vår språkbruk ytterligere.

Takk til Jan Odgaard-Jensen for bistand til statistiske analyser.

REFERANSER

- 1 Glenton C, Santesso N, Rosenbaum S, Nilsen ES, Rader T, Ciapponi A, et al. Presenting the results of Cochrane Systematic Reviews to a consumer audience: a qualitative study. *Med Decis Making* 2010; **30**: 566-77.
- 2 Akl EA, Oxman AD, Herrin J, Vist GE, Terrenato I, Sperati F, et al. Using alternative statistical formats for presenting risks and risk reductions. *Cochrane Database Syst Rev* 2011; (3): CD006776.
- 3 Spiegelhalter D, Pearson M, Short I. Visualizing uncertainty about the future. *Science* 2011; **333**: 1393-400.
- 4 Trevena LJ, Davey HM, Barratt A, Butow P, Caldwell P. A systematic review on communicating with patients about evidence. *J Eval Clin Pract* 2006; **12**: 13-23.
- 5 Rosenbaum SE, Glenton C, Nylund HK, Oxman AD. User testing and stakeholder feedback contributed to the development of understandable and useful Summary of Findings tables for Cochrane reviews. *J Clin Epidemiol* 2010; **63**: 607-19.
- 6 Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011; **64**: 401-6.
- 7 Nasjonalt kunnskapssenter for helsetjenesten. Slik oppsummerer vi forskning. Håndbok for Nasjonalt kunnskapssenter for helsetjenesten. 2013.
- 8 Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011; **64**: 383-94.
- 9 Ezzo J, Bausell B, Moerman DE, Bertram B, Hadhazy V. Reviewing the reviews. How strong is the evidence? How clear are the conclusions? *Int J Technol Assess Health Care* 2001; **17**: 457-66.