
Nivået på dokumentasjonen og styrken av anbefalingene i kliniske retningslinjer

KRONIKK

ANDREW D. OXMAN

SIGNE FLOTTORP

Seksjon for helsetjenesteforskning
Statens institutt for folkehelse
Postboks 4404 Torshov
0403 Oslo

JOHN G. COOPER

4130 Hjelmeland

PER HJORTDAHL

Institutt for allmennmedisin og samfunnsmedisinske fag
Universitetet i Oslo
Postboks 1130 Blindern
0317 Oslo

SVERRE SANDBERG

Laboratorium for klinisk biokjemi
Haukeland Sykehus
5021 Bergen

LARS H. VORLAND

Mikrobiologisk avdeling
9038 Regionsykehuset i Tromsø

I denne utgaven av Tidsskriftet presenterer vi retningslinjer for diagnostikk og behandling av vannlatingsplager hos kvinner og av sår hals. I denne artikkelen beskriver vi tilnærmingen vi brukte for å vurdere ”nivået på dokumentasjonen” og ”styrken av anbefalingene” i disse retningslinjene. Vi definerte ”nivået på dokumentasjonen” som den grad av sikkerhet vi kan ha for at estimatet av en effekt eller assosiasjon er korrekt. Med ”styrken av anbefalingene” mener vi den grad av sikkerhet vi kan ha for at det å følge anbefalingene vil gjøre mer nytte enn skade.

Vi valgte en enkel tilnærming som kunne brukes for å vurdere dokumentasjon om nøyaktighet av diagnostiske tester, insidens eller prevalens, så vel som bivirkninger og effekt av behandlingstiltak. Tilnærmingen er primært basert på studiedesign. Nivå 1-dokumentasjon indikerer at vi fant en god systematisk oversikt med minst én studie, nivå 2 at vi fant minst én god studie, og nivå 3 at vi ikke fant gode studier. Hva som vurderes som en ”god systematisk oversikt” og en ”god studie” blir definert ved å bruke eksplisitte kriterier. Styrken av anbefalingene er basert på nivået på dokumentasjonen, dessuten vurdering av appliserbarhet, avveininger og verdier.

Vi drøfter fordeler og ulemper ved vår tilnærming sammenliknet med andre tilnærminger.

Bør du gi penicillin til en 25 år gammel mann med sår hals? Bør du ta en dyrkingsprøve fra halsen for å diagnostisere β -hemolytisk streptokokktonsillitt først? Eller bør du bruke en hurtigtest? Bør du behandle en 35 år gammel kvinne med svie ved vannlating og hyppig vannlating med antibiotika? Hvis ja – i hvor mange dager? Bør urinen dyrkes? Bør urinen undersøkes mikroskopisk? Bør du benytte urinstrimmel for å undersøke urinen?

Klinikere blir daglig konfrontert med slike problemer. Hvilke spørsmål skal stilles til pasientene, hva skal den kliniske undersøkelsen omfatte, hvilke diagnostiske tester skal rekvireres, hvilken informasjon skal pasienten få, hvilke tiltak skal anbefales og hva slags oppfølging er nødvendig? Som støtte for slike avgjørelser må legene, bevisst eller ubevisst, benytte seg av retningslinjer – enkle beslutningsregler for å løse sammensatte problemer (1), som for eksempel: ”Pasienter med sår hals skal vanligvis behandles symptomatisk uten antibiotika” (2).

Kvaliteten av behandlingen vil avhenge av kvaliteten av retningslinjene som benyttes. I denne utgaven av Tidsskriftet presenterer vi retningslinjer for diagnostikk og behandling av sår hals og av akutte vannlatingsplager hos kvinner (2, 3). Før retningslinjer tas i bruk bør klinikerne spørre seg: Hva er dokumentasjonen som underbygger disse retningslinjene og hvor sikkert er det at det vil føre til mer nytte enn skade å følge dem? Mer detaljerte spørsmål for å veilede i kritisk vurdering av kliniske retningslinjer (tab 1) er beskrevet andre steder (4 – 8).

Tabell

Tabell 1 Kritiske spørsmål til kliniske retningslinjer (5, 6, 8)
<i>Er det sant? Er retningslinjene til å stole på?</i>
Er alle viktige beslutninger og utfall spesifisert?
Ble en eksplisitt og pålitelig prosess brukt for å identifisere, velge ut og sammenstille de forskningsresultater som er relevante for hver beslutning?

Tabell 1 Kritiske spørsmål til kliniske retningslinjer (5, 6, 8)
Ble en eksplisitt og pålitelig prosess brukt for å belyse verdien av de forskjellige utfall?
Er retningslinjene basert på oppdatert kunnskap?
Kan usikkerhet knyttet til underlagsmaterialet, de effektvurderinger eller den verdifastsettelsen man bygger på, føre til endringer av anbefalingene?
Er dokumentasjonen og anbefalingene gradert?
<i>Hva er resultatene? Er retningslinjene viktige?</i>
Er det stor variasjon i nåværende praksis?
Er det gitt praktisk gjennomførbare og klinisk viktige anbefalinger?
Omfatter anbefalingene så mange mennesker eller mennesker med så høy risiko for sykdom eller en så ressurskrevende praksis at endringer i hvordan problemet håndteres virkelig vil kunne bety en forskjell?
<i>Vil resultatene hjelpe meg i min praksis?</i>
Er formålet med retningslinjene overensstemmende med hva du prøver å oppnå i din kliniske praksis?
Er anbefalingene akseptable for dine pasienter?
Hvilke hindringer foreligger som vil gjøre det vanskelig å implementere retningslinjene, og kan de overvinnes?

Vår tilnærming

I denne artikkelen beskriver vi en tilnærming som vi brukte for å vurdere dokumentasjonen for våre anbefalinger og hvor sikre vi mener vi kan være for at det vil føre til mer nytte enn skade å følge retningslinjene. Vi sammenlikner vår tilnærming med andres og diskuterer fordeler og ulemper ved de ulike tilnærmingene. Det er viktig at de som utvikler retningslinjer, veiledere, metodebøker og liknende, bruker en systematisk og eksplisitt tilnærming til dette, og at brukerne forstår hva som ligger bak slike anbefalinger.

Hva menes med ”nivået på dokumentasjonen” og ”styrken av anbefalingene”?

Grunnlaget for kliniske retningslinjer har ofte vært implisitt, usystematisk og basert på ekspertenes meninger. Det har ofte manglet klar referanse til dokumentasjon, eller det har manglet dokumentasjon for disse meningene. Svakheter ved denne måten å utarbeide retningslinjer på har etter hvert fått økende oppmerksomhet. I Norge og internasjonalt har det vært nedlagt mye arbeid for å ta i bruk mer systematiske og eksplisitte metoder (4 – 8). Som en del av denne innsatsen er det utviklet eksplisitte kriterier for å karakterisere ”nivået på dokumentasjonen” som underbygger kliniske retningslinjer og ”styrken av anbefalingene”. Dette skal hjelpe brukeren til å finne ut av i hvilken grad man kan stole på resultatene i de studiene som retningslinjene bygger på,

og i hvilken grad man kan ha tillit til at det er riktig å følge anbefalingene. I tabell 2 (9) og tabell 3 har vi gjort rede for hva vi mener med ”nivået på dokumentasjonen” og ”styrken av anbefalingene”.

Tabell

Tabell 2 Hva er ”nivået på dokumentasjonen”?
”Nivået på dokumentasjonen” er den grad av sikkerhet vi kan ha for at estimatet av en effekt eller en assosiasjon er korrekt. For estimer av effekt tilsvarer dette styrken på en slutning om årsakssammenheng. De følgende spørsmålene er da relevante (9):
– Hvor valide er de inkluderte studiene?
– Hvor store og presise er de observerte assosiasjonene?
– Hvor konsistente er assosiasjonene på tvers av studiene?
– Er det en tydelig sammenheng mellom dose og respons?
– Er det indirekte dokumentasjon som underbygger slutningen?
– Er andre plausible konkurrerende forklaringer på den observerte assosiasjonen blitt utelukket?

Tabell

Tabell 3 Hva er ”styrken av anbefalingene”?
”Styrken av anbefalingene” er den grad av sikkerhet vi kan ha for at det å følge anbefalingen vil gjøre mer nytte enn skade. Styrken av en anbefaling avhenger av informasjon om og vurdering av:
– Nivået på dokumentasjonen
– Appliserbarhet av dokumentasjonen
– Avveininger mellom forventede positive virkninger, skadevirkninger og kostnader

Vår tilnærming til gradering av nivået på dokumentasjonen og styrken på anbefalingene

Vi valgte en enkel tilnærming som kan brukes for å vurdere dokumentasjon både om nøyaktighet av diagnostiske tester, insidens eller prevalens så vel som bivirkninger og effekt av behandlingstiltak (10 – 13). Ulike spørsmål krever ulike metoder for at vi skal ha tillit til dokumentasjonen. Vårt mål var å finne frem til en tilnærming som tydelig reflekterer i hvilken grad vi føler oss trygge på at det å følge anbefalingene, vil føre til de resultater som kan forventes, basert på den beste dokumentasjonen som er tilgjengelig.

Tilnærmingen som vi har brukt (tab 4) er primært basert på vurdering av metodene som er benyttet i studiene, med følgende tillegg:

Tabell

Tabell 4 Nivået på dokumentasjonen og styrken av anbefalingene		
Nivået på dokumentasjonen		
1	Meget god	Vi fant en god systematisk oversikt med minst én god studie
2	God	Vi fant minst én god studie
3	Manglende	Vi fant ingen gode studier
Det er brukt eksplisitte kriterier for å vurdere kvaliteten av oversikter og studier (tab 5)		
Styrken av anbefalingene		
A	Sterk	Basert på meget god dokumentasjon (nivå 1) og enighet i referansegruppen om validiteten av dokumentasjonen, appliserbarhet og avveining av nytte og ulempe
B	Moderat	Basert på minst én god studie (nivå 2) og enighet om validiteten av dokumentasjonen, appliserbarhet og avveining av nytte og ulempe
C	Svak	Basert på manglende dokumentasjon (nivå 3) og enighet om appliserbarhet og avveining av nytte og ulempe
Styrken på anbefalingene var primært basert på nivået på dokumentasjonen. Vi tok også hensyn til appliserbarhet av dokumentasjonen i norsk allmennpraksis i dag. På grunnlag av tilgjengelig dokumentasjon veide vi forventede fordeler mot forventede ulemper og kostnader ved de alternativer som ble vurdert. Styrken av anbefalingene reflekterer også i hvilken grad det var enighet om disse vurderingene i referansegruppen som utviklet retningslinjene og blant dem som kommenterte utkastet til retningslinjer. En eksplisitt forklaring ble gitt når styrken på anbefalingene ikke samsvarte direkte med nivået på dokumentasjonen		

- Nivå 1-dokumentasjon indikerer at det finnes en god systematisk oversikt, at oversikten inkluderer studier med god design, og at det ikke er uenighet om at dokumentasjonen er god.
- Hva som vurderes som en ”god systematisk oversikt” og en ”god studie” blir definert ved å bruke eksplisitte kriterier (tab 5).

Tabell

Tabell 5 Inklusjonskriterier og kriterier for vurdering av kvaliteten av studiene ved gradering av nivået på dokumentasjonen (10 – 13)			
Inklusjonskriterier	Møtt	Tvil	Ikkemøtt
<i>For systematiske oversikter:</i>			
Oversikten må inneholde en metodeseksjon som beskriver inklusjonskriteriene og søkestrategien	å å	å	å å
<i>For studier av effekt :</i>			
Fordelingen av pasienter til behandlingene må være randomisert	å	å	å å
<i>For studier av nøyaktighet av diagnostiske tester :</i>			
Det må være en uavhengig, blind sammenlikning med en referansestandard (gullstandard) for diagnosen	å	å	å å

Tabell 5 Inklusjonskriterier og kriterier for vurdering av kvaliteten av studiene ved gradering av nivået på dokumentasjonen (10 – 13)			
Den diagnostiske testen må ha blitt evaluert i et relevant spekter av pasienter (tilsvarende den gruppen som testen vil bli brukt på i praksis)	å	å	å å
<i>For studier av skade eller bivirkninger :</i>			
Det må være en klart identifisert sammenlikningsgruppe for dem som har økt risiko for, eller som har det utfallet som er av interesse (enten fra randomiserte, kvasirandomiserte eller ikke-randomiserte, kontrollerte forsøk, kohortstudier med statistiske justeringer for å skape sammenliknbare grupper eller case control-studier)	å	å	å å
Eksposering for intervensjonen og de kliniske utfall må være målt på samme måten i begge grupper	å	å	å å
Oppfølgingstiden må være lang nok	å	å	å å
<i>For studier av insidens eller prevalens :</i>			
Det må være et representativt (tilfeldig, konsekutivt eller totalt) utvalg	å	å	å å
Det må være klart definerte diagnostiske kriterier	å	å	å å
Videre vurdering kun av inkluderte studier	å	å	å å
<i>For systematiske oversikter :</i>			
Var kriterier for vurdering av validitet av inkluderte studier og analysene beskrevet i metodeseksjonen?	å	å	å å
Var det inkludert ³ 1 studie som møter inklusjonskriteriene for enkeltstudier?	å	å	å å
Var resultatene av de inkluderte studiene som er relevante for anbefalingene i hovedsak konsistente?	å	å	å å
<i>For studier av effekt :</i>			
Var randomiseringsprosessen lukket?	å	å	å å
Var sammenlikningsgruppene like i utgangspunktet?	å	å	å å
Var ³ 80 % av pasientene som ble inkludert i forsøket redegjort for i konklusjonen?	å	å	å å
Var studien dobbeltblindet?	å	å	å å
<i>For studier av nøyaktighet av diagnostiske tester :</i>			
Var referansestandard applisert uavhengig av resultatet av den diagnostiske testen?	å	å	å å
Hvis studien gjaldt et diagnostisk kriteriesett eller algoritme, ble dette validert i en annen kohort av pasienter?	å	å	å å
<i>For studier av skade eller bivirkninger :</i>			
Var sammenlikningsgruppene like på andre viktige måter enn eksponeringen for intervensjonen?	å	å	å å

Tabell 5 Inklusjonskriterier og kriterier for vurdering av kvaliteten av studiene ved gradering av nivået på dokumentasjonen (10 – 13)			
Var ³ 80 % av pasientene som ble inkludert i studien redegjort for i konklusjonen?	å	å	å å
Var det brukt blindet vurdering eller objektive mål for å måle eksponering eller utfall?	å	å	å å
<i>For studier om insidens eller prevalens :</i>			
Var de diagnostiske kriterier applisert på samme måten til hele utvalget slik at det var liten risiko for falskt positive eller falskt negative tilfeller?	å	å	å å

– Andre faktorer enn de som er listet opp i tabell 4 og tabell 5, kan påvirke tilliten vi har til dokumentasjonen. Vi har ikke forsøkt å utvikle regler for når nivået på dokumentasjonen burde modifiseres, men vi har valgt å være eksplisitte i forhold til om det var enighet eller ikke og til å oppgi våre grunner når vi justerte nivået på dokumentasjonen.

Tilnærmingen vi brukte for å gradere styrken på anbefalingene, var også først og fremst basert på nivået på dokumentasjonen, sammen med en vurdering av i hvilken grad det var enighet om appliserbarheten av dokumentasjonen og avveiningen eller verdien av de forskjellige utfallene. Vi brukte vårt eget skjønn for å ta stilling til om fordelene som kunne forventes hvis en anbefaling ble fulgt, ville være verd de forventede skadelige virkningene og kostnadene. Hvis pasienter kunne komme til forskjellige konklusjoner på grunn av ulike verdier knyttet til disse avveiningene, anbefalte vi at de blir gitt relevant informasjon og mulighet til å ta informerte valg på grunnlag av sine egne verdier og preferanser. Vi testet vårt eget skjønn ved at kolleger ble spurt om kritisk å gjennomgå utkastet til retningslinjer og til spesifikt å kommentere våre avveininger mellom fordeler og ulemper ved anbefalingene. Vi drøftet også disse avveiningene i fokusgrupper med medarbeidere og med pasienter. Hvis styrken på anbefalingen ikke korresponderte direkte med nivået på dokumentasjonen, oppgav vi årsaken til dette.

Diskusjon

Canadian Task Force on the Periodic Health Examination gjorde i 1979 et av de første forsøk på eksplisitt å karakterisere nivået på dokumentasjon som underbygde anbefalingene om helsetiltak, og styrken av disse anbefalingene (14). Siden den gang er en rekke alternative metoder for å klassifisere kliniske retningslinjer blitt foreslått og benyttet (15 – 23).

Tilnærmingen som ble benyttet av Canadian Task Force, er basert på studienes design. Randomiserte, kontrollerte forsøk ble klassifisert som god dokumentasjon, kohort- og case control-studier ble klassifisert som brukbar dokumentasjon og ekspertmeninger alene som manglende dokumentasjon. Det er en direkte sammenheng mellom styrken på anbefalingen og nivået på dokumentasjonen, slik at en sterk anbefaling korresponderer med at det er god dokumentasjon. Denne metoden er relativt enkel og derfor tiltalende. Den skiller mellom nivået på dokumentasjonen og styrken av

anbefalingene, samtidig som det siste er knyttet til det første. Fordi metoden er enkel, er den lett å forstå, lett å bruke og lett å presentere. Dette systemet kan imidlertid bare brukes til å vurdere dokumentasjon av effekten av tiltak. Det omfatter ikke vurdering av dokumentasjon som krever andre studiedesigner enn randomiserte forsøk. Systemet er dessuten så forenklet at mange avgjørelser må gjøres implisitt. For eksempel:

- Bør små og dårlige randomiserte, kontrollerte forsøk betraktes som nivå 1-dokumentasjon?
- Hvis det er betydningsfull variasjon (heterogenitet) i resultatene av randomiserte, kontrollerte forsøk, bør dette fortsatt betraktes som nivå 1-dokumentasjon?
- Bør overbevisende resultater fra ikke-eksperimentelle studier alltid oppfattes som nivå 2-dokumentasjon?

Denne tilnærmingen krever også implisitte avgjørelser om styrken av anbefalingene, for eksempel hvordan forventede virkninger med hensyn til nytte, skade og kostnader blir avveid og tatt hensyn til når man går fra en vurdering av hvor solid dokumentasjonen er til hvilke følger resultatene har for praksis.

Senere tilnærminger basert på studiedesign har føyd til flere detaljer, for eksempel ved å skille mellom store og vel utførte randomiserte, kontrollerte forsøk og små randomiserte, kontrollerte forsøk, men de har fortsatt sine begrensninger (15, 16). Fremgangsmåter basert på systematiske oversikter har fordeler i forhold til fremgangsmåter basert på studiedesign ved eksplisitt å gjøre rede for hvordan dokumentasjonen er identifisert, valgt ut, kritisk vurdert og summert. Men dette blir mer komplisert, og systemet lider fortsatt under at man forsøker å foreta et sammensatt sett med vurderinger ved å bruke et altfor enkelt sett med regler (17). Noen fremgangsmåter klassifiserer systematiske oversikter eller metaanalyser som nivå 1-dokumentasjon, uten å ta hensyn til at en systematisk oversikt kan avdekke hva som helst, fra fullstendig mangel på dokumentasjon til svært solid dokumentasjon (18 – 21). Disse tilnærmingene beskriver heller ikke hvordan man går fra vurderinger av nivået på dokumentasjonen til vurderinger av implikasjonene av denne dokumentasjonen.

I det siste er det blitt argumentert for en tilnærming basert på systematiske oversikter som eksplisitt tar hensyn til klinisk betydning, statistisk signifikans og basisrisiko i tillegg til studiedesign (22). Denne metoden taper mer av enkelheten i forhold til den opprinnelige tilnærmingen til The Canadian Task Force. En annen ulempe er at vurderingen av nivået på dokumentasjonen ikke er atskilt fra vurderingen av styrken av anbefalingene. Dette systemet lider i enda større grad av forsøket på å redusere komplekse vurderinger til en enkel modell. Slike enkle modeller tar ikke hensyn til alle de forskjellige faktorene som avgjør hvilken tillit vi har til dokumentasjonen, eller anbefalingene som utledes fra dokumentasjonen. Dramatisk store effekter som er konsistente på tvers av ikke-eksperimentelle studier kan for eksempel være så overbevisende at det ville være uetisk å utføre et randomisert kontrollert forsøk.

En annen fremgangsmåte som er blitt brukt for å karakterisere nivået på dokumentasjonen, er basert på kriterier for å vurdere styrken på en slutning om kausalitet (23), slik som sjekklisten til Bradford-Hill (24) eller spørsmålene i tabell 2. Denne tilnærmingen er begrepsmessig tiltalende, fordi den får tak i det essensielle når det gjelder hva slutninger om ”nivået på dokumentasjonen” handler om. I likhet med andre tilnærminger basert på systematiske oversikter svekkes den imidlertid av vanskelighetene med å forenkle sammensatte vurderinger til enkle modeller.

Vår fremgangsmåte er ikke ideell. Argumentene for at vi har gått frem som beskrevet, var at tilnærmingsmåten var relativt enkel, som The Canadian Task Force og liknende metoder, mens den samtidig

- – inkorporerte en mer eksplisitt fremgangsmåte for å avgjøre hvordan ”god” var definert
- – tok hensyn til behovet for en systematisk oversikt som grunnlag for vurdering av nivået på dokumentasjonen
- – gjorde eksplisitte de vurderinger som førte til justeringer i nivået på dokumentasjonen eller styrken av en anbefaling
- – unngikk i det minste noen av fallgruvene som oppstår når man forsøker å bygge opp en enkel modell for å fange et sammensatt sett med vurderinger
- Noen svakheter ved vår tilnærming er:
 - – Kriteriene for kritisk vurdering av studier omfatter ikke alle faktorer som kan påvirke tilliten til dokumentasjonen.
 - – Tilnærmingen gir ikke veiledning om hvordan graderingen av nivået på dokumentasjonen og styrken på anbefalingene bør justeres når det er svakheter ved studiene.
 - – Graderingen av styrken på anbefalingene innebærer komplekse og delvis skjønnsmessige vurderinger som det er vanskelig å gjøre fullstendig rede for.

Konklusjon

Vi har utviklet en metode for å gradere nivået på dokumentasjonen og styrken av anbefalingene i kliniske retningslinjer. Det er behov for videre arbeid for å utvikle en fornuftig fremgangsmåte som kan anvendes for alle de ulike typer av dokumentasjon som må ligge under anbefalinger om helsetjenester. Metoden bør kunne være i alminnelig bruk og være lett å forstå for en større gruppe brukere, inkludert klinikere, pasienter og beslutningstakere. Det kan argumenteres for og imot de for tiden tilgjengelige metoder. Det at det benyttes så mange forskjellige metoder, avleder oppmerksomheten fra en av de viktigste grunner for eksplisitt å gradere nivået på dokumentasjonen og styrken på anbefalingene – så konsist som mulig å karakterisere og formidle disse to viktige aspektene ved anbefalinger om helsetjenester. Hvis klinikere og andre beslutningstakere som skal bruke anbefalinger om helsetiltak, kan vurdere hvor solid dokumentasjonen bak anbefalingene er og hvor sterke anbefalingene er, blir det lettere å ta informerte og gode beslutninger.

LITTERATUR

1. Eddy DM. Clinical policies and the quality of clinical practice. N Engl J Med 1982; 307: 343 – 7.
2. Flottorp S, Oxman AD, Cooper JG, Hjortdahl P, Sandberg S, Vorland LH. Retningslinjer for diagnostikk og behandling av sår hals Tidsskr Nor Lægeforen

2000; 120: 1754 – 60.

3. Flottorp S, Oxman AD, Cooper JG, Hjortdahl P, Sandberg S, Vorland LH. Retningslinjer for diagnostikk og behandling av akutte vannlatingsplager hos kvinner Tidsskr Nor Lægeforen 2000; 120: 1748 – 53.
4. Retningslinjer for retningslinjer. Prosess og metode for utvikling og implementering av faglige retningslinjer. Oslo: Statens helsetilsyn, 1998.
5. Hayward RSA, Wilson MC, Tunis SR, Bass EB, Guyatt GH, for the Evidence-Based Working Group. Users' guides to the medical literature, VIII: How to use clinical practice guidelines, A: are the recommendations valid? JAMA 1994; 274: 570 – 4.
6. Wilson MC, Hayward RSA, Tunis SR, Bass EB, Guyatt GH, for the Evidence-Based Working Group. Users' guides to the medical literature, VIII: How to use clinical practice guidelines, B: What are the recommendations and will they help you in caring for your patients? JAMA 1994; 274: 1630 – 2.
7. Sackett DL, Richardson S, Rosenberg W, Haynes RB, Straus S. Evidence-based medicine: how to practice & teach EBM. 2. utg. London: Churchill Livingstone, 1999.
8. Bjørndal A, Flottorp S, Klovning A. Medisinsk kunnskapshåndtering. Oslo: Gyldendal Akademisk, 2000.
9. Clarke M, Oxman AD, red. Strength of evidence. Cochrane Reviewers' Handbook. 4. utg. 4.0. [oppdatert juli 1999]. I: The Cochrane Library, Issue 1. Oxford: Update Software, 2000.
10. Jaeschke R, Guyatt GH, Sackett DL, for the Evidence-Based Working Group. Users' guides to the medical literature, III: How to use an article about a diagnostic test, A: Are the results of the study valid? JAMA 1994; 271: 389 – 91.
11. Levine M, Walter S, Lee H, Haines T, Holbrook A, Moyer V, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, IV: How to use an article about harm. JAMA 1994; 271: 1615 – 9.
12. Richardson WS, Wilson MC, Guyatt GH, Cook DJ, Nishikawa J, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, XV: How to use an article about disease probability for differential diagnosis? JAMA 1999; 281: 1214 – 9.
13. Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Working Group. Users' guides to the medical literature, II: How to use an article about therapy or prevention, A: Are the results of the study valid? JAMA 1993; 270: 2598 – 601.
14. Canadian Task Force on the Periodic Health Examination. The periodic health examination. Can Med Assoc J 1979; 121: 1193 – 254.
15. Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. Chest 1986; 89 (suppl 2): 2 – 3.

16. US Preventive Services Task Force. Guide to Clinical Preventive Services. 2. utg. Baltimore: Williams & Wilkins, 1996; xxxix-lv.
 17. Ball C, Sackett D, Phillips B, Straus S, Haynes B. Levels of evidence and grades of recommendations. Last revised 17 September 1998. Centre for Evidence-Based Medicine. <http://cebm.jr2.ox.ac.uk/docs/levels.html> (22.9.1999).
 18. Cook DJ, Guyatt GH, Laupacis A, Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. Antithrombotic Therapy Consensus Conference. Chest 1992; 102 (suppl 4): 305 – 11.
 19. US Department of Health and Human Services, Public Health Service, Agency Health Care Policy and Research (1992). Acute pain management: operative or medical procedures and trauma. AHCPR Pub 92 – 0038. Rockville, MD: Agency for Health Care Policy and Research Publications, 1992.
 20. Petrie J, Barnwell E, Grimshaw J. Criteria for appraisal for national use. Pilot Edition, September 1995. Scottish Intercollegiate Guidelines Network. <http://show.cee.hw.ac.uk/sign/criteria.htm> (22.9.1999).
 21. Eccles M, Clapp Z, Grimshaw J, Adams PC, Higgins B, Purves I et al. North of England evidence based guidelines development project: methods of guideline development. BMJ 1996; 312: 760 – 2.
 22. Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. IX. A method for grading health care recommendations. JAMA 1995; 274: 1800 – 4.
 23. Gyorkos TW, Tannenbaum TN, Abrahamowicz M, Oxman AD, Scott EA, Millson ME et al. An approach to the development of practice guidelines for community health interventions. Can J Public Health 1994; 85 (suppl 1): 8 – 13.
 24. Hill AB. Principles of medical statistics. 9. utg. London: Lancet, 1971: 312 – 20.
-

Publisert: 10. juni 2000. Tidsskr Nor Legeforen.

Opphavsrett: © Tidsskriftet 2026 Lastet ned fra tidsskriftet.no 2. august 2026.