

# Evaluering av kvalitativ forskning – er det ønskjeleg og mogeleg med ei ny tilnærming?

*Michael 2010; suppl 9: 61–68.*

*Som forskar, fagfelle og tidsskriftsredaktør har eg gjort erfaringar som har utvikla seg til ei vurdering som eg nok deler med mange: Der er nokre svært krevjande utfordringar knytt til evaluering av kvalitativ forskning. Litteratur på feltet peikar ofte på skilnader i kunnskapssyn og forskningstradisjonar. Eit mangfald av perspektiv på omgrep som sanning og metode gjer det ikkje lett å semjast om kriterium for evaluering. I denne artikkelen gjer eg først greie for nokre av dei erfaringane som har fått meg til å sjå på evaluering av kvalitativ forskning som ei akademisk utfordring. Med referanse til ein nyleg publisert artikkel av Stige, Malterud og Midtgarden, skisserer eg så opp tre grupper av løysingsforslag. Ingen av desse vert vurderte som fullt ut tilfredsstillande og det vert argumentert for at kvalitativ forskning er truga av ein fare for fragmentert evalueringspraksis, vag evalueringspraksis og/eller evaluering på feil premiss. Regel- og kriteriebasert evaluering vert difor foreslått erstatta av ei agendabasert tilnærming med vekt på dialog og refleksivitet og med meir rom for pluralisme. Som reiskap for løysingsorienterte evalueringsamtalar vert så ein konkret agenda lansert (EPICURE-Engagement, Processing, Interpretation, Critique, Usefulness, Relevance og Ethics). EPICURE har ein dimensjon som viser til forskinga si utfordring i høve til å utforme truverdige og engasjerande framstillingar av dei fenomenane som vert studerte, og ein dimensjon som viser til forskinga sin relasjon til eit praksisfelt. I ei avsluttande oppsummering drøftar eg implikasjonane av at vi enno ikkje heilt veit korleis ein agendabasert og dialogisk orientert evalueringspraksis best kan utformast slik at han fungerer både for forfattarar, fagfellar, redaktørar og lesarar. Dette gjenstår å prøve ut, evaluere og forske på.*

I 2001 publiserte Kirsti Malterud to artiklar om kvalitativ forskning i tidsskriftet *The Lancet* (1,2). Desse artiklane er både kjende og anerkjende, ikkje minst i høve til temaet evaluering av kvalitativ forskning. Det er eit

spørsmål som eg sjølv har vore opptatt av og absolutt var fokusert på i 2001. Det tok likevel litt tid før eg oppdaga artiklane til Malterud. Litt synd, tenkjer eg i dag. Litt underleg kanskje og. Men, som musikkterapeut er eg vel ikkje av dei mest flittige lesarane av *The Lancet* og i dei søka etter litteratur som eg gjorde på den tida var eg ikkje medviten nok på at eg som musikkterapiforskar kan møte på mange av dei same metodeutfordringane som ein forskar i allmennmedisin. Eg henta inspirasjon hjå kvalitative forskarar i musikkterapi, psykologi og samfunnsfag og hadde ikkje heilt tatt inn at ei spennande utvikling på dette feltet var på gang også innan medisinen. I dag kan eg sjå tilbake på eit givande samarbeid med Kirsti Malterud, nettopp med fokus på evaluering av kvalitativ forskning (3). Før eg fortel nærare om resultatet av dette samarbeidet og reflekterer over moglege implikasjonar, vil eg gjere greie for dei faglege erfaringane som har fått meg til å leggje vekt på evaluering av kvalitativ forskning som ei viktig akademisk utfordring av stor relevans for ulike helserelevante fag.

Ein av grunnane til at eg i 2001 var opptatt av spørsmålet om evaluering av kvalitativ forskning var at eg ni år tidlegare hadde tatt initiativ til etablering av *Nordic Journal of Music Therapy*. Dette er eit fagfellevurdert tidsskrift som i og for seg har greidd seg bra. Det er eit nivå 2-tidsskrift i dag, med publiseringssavtale med Routledge og med eit internasjonalt, tverrfagleg nedslagsfelt. Men som dåverande redaktør sleit eg med ei utfordring som eg opplevde som vanskeleg: Korleis sikre god og grundig evaluering av kvalitative artiklar? Det var på ein måte litt overraskande for meg at dette skulle vere så krevjande. Den nordiske musikkterapi tradisjonen hadde sin styrkje nettopp i kvalitative metodar og min eigen kompetanse som redaktør var også vel så god innan kvalitativ som innan kvantitativ forskning. Likevel erfarte eg gong på gong at det var enklare å få i hamn gode evalueringssprossar av kvantitative enn av kvalitative artiklar.

Desse vanskanane handla mellom anna om uvanleg store sprik i fagfellane sine vurderingar når det kom til kvalitative artiklar. Det som av ein fagfelle vart oppfatta som spennande nybrottsarbeid, kunne av ein annan verte vurdert som for laust eller vidløftig. Det som av ein fagfelle vart oppfatta som tamt og gammalt nytt, kunne av ein annan verte vurdert som solid og godt handverk. Sprik i vurderingar skulle i utgangspunktet likevel vere handterleg, dels ved å hente inn ei tredje eller fjerde vurdering, dels ved å drøfte premiss for og implikasjonar av dei ulike vurderingane. Det siste var likevel verre enn venta. Ikkje alle fagfellar var innstilte på å gjere greie for bakgrunnen for eigne vurderingar. Ikkje alle forfattarar var så innstilte på dialog heller. Dei var kanskje meir opptatt av at fagfellen ikkje hadde forstått.

Slike problem kan beskrivast som manglande evne og vilje til metakritikk (4). Eg må innrømme at eg dei første åra i redaktørstolen i *Nordic Journal of Music Therapy*, altså tidleg på nittitalet, tenkte at eg først og fremst sto over for utfordringar knytt til at musikkterapifaget framleis var relativt ungt som akademisk felt. I meir etablerte fagområde var det sikkert annleis. Eg skulle ikkje studere metodelitteraturen i andre fag særleg nøye før eg såg at temaet var meir allment enn som så. Der føregjekk og det føregår mange debattar om korleis evaluere kvalitativ forskning. Dei fleste debattantane peikar på utfordringar knytt til at der innan kvalitativ forskning er ulike paradigme og tradisjonar med temmeleg ulike perspektiv på forholdet mellom språk og røyndom. Med pluralisme i oppfatningar av omgrep som kunnskap, sanning og metode er det ikkje lett å semjast om kriterium for evaluering. I 2002 prøvde eg å gje perspektiv til debatten i eige fag ved å skrive ein diskusjonsartikkel (5) der eg foreslo at Wittgensteins metaforar om språkspel og familielikskap (6) kunne vere til hjelp, både for å forstå og å handtere utfordringane. Dersom vi samanliknar dei ulike forskningstradisjonane med ulike spel, til dømes ballspel, ser vi at kriteriebasert vurdering i kvalitativ forskning vil kunne vere like forvirrande og frustrerande som å dømme ein fotballkamp etter handballreglar. Så lar kriterieproblemet seg kanskje løyse ved å finne fram det rette settet med reglar, til dømes ved å innføre ulike sjekklister for ulike tradisjonar innan kvalitativ forskning? Eg trur ikkje det, mellom anna av di forskning som praksis har ein annan karakter enn dei spela eg refererte til over. Forsking inviterer ikkje berre til kreativitet innanfor gjeldande reglar, men også til kreativitet som sprenger eksisterande rammer.

Den løysinga eg foreslo i 2002 – altså utan å kjenne til Malterud sitt arbeid – var ikkje så ulik den løysinga ho foreslo i *The Lancet* året før, nemleg å artikulere metakriterium eller generelle standardar som kunne akseptierast på tvers av ulike forskningstradisjonar. I dag oppfattar eg dette som ei ufullstendig løysing. Eg vil gjere greie for kvifor og også skissere eit nytt forslag til løysing, ved å referere til ein artikkel eg nyleg har skrive saman med Kirsti Malterud og Torjus Midtgarden. Forslaget byggjer på det premiss at refleksivitet er grunnleggjande i kvalitativ forskning, forstått som vilje og evne til kritisk å undersøkje eigen posisjon og eigne perspektiv og føresetnader som deltakar i forskingsprosessen. Ein slik kritisk refleksjon føreset dialog og samarbeid med andre aktørar i prosessen (7,8).

### **Steg i retning av ei agendabasert tilnærming til evaluering**

Artikkelen «Toward an agenda for evaluation of qualitative research» (3) kan karakteriserast som teoretisk forskning. Det materialet som vert undersøkt

er eksisterande tekstar på feltet og målet med artikkelen er å analysere og vurdere desse og eventuelt artikulere alternative omgrep og tilnærmingar, dersom ein ikkje finn eksisterande forslag tilfredsstillande (9). Forfattarane sin erfaringsbakgrunn er då ein integrert del av analysen og forfattargruppa si samansetning er slik sett ein del av metoden. Med ein forskar innan musikkterapi, ein innan medisin og ein innan vitskapsteori har vi prøvd å sikre oss breidde i vurderingsgrunnlaget, både i høve til praktiske, teoretiske og metateoretiske kontekstar. Eg skal i det følgjande kort gjere greie for nokre hovudmoment i argumentasjonen utvikla i denne artikkelen, der eg då vil presisere at dette samandraget får stå for mi eiga rekning:

Gjennomgang av litteraturen gjev eit bilete av grovt sett tre ulike løysingsforslag for kriterieproblemet i kvalitativ forskning: Somme forfattarar argumenterer for bruk av kriterium relative til aktuelt prosjekt eller tradisjon (lokale kriterium), somme argumenterer for generelle standardar (metakriterium) som kan brukast på tvers av ulike tradisjonar, medan somme argumenterer for at ålmenne, spesifikke kriterium (til dømes i form av sjekklister) likevel lar seg artikulere, trass i ulike tradisjonar. Ingen av desse løysingane er fullt ut tilfredsstillande. Lokale kriterium er vanskeleg å bruke i praksis, då det ikkje alltid er klårt kva kriterium som skal gjelde, mellom anna av di mange forskingsprosjekt hentar impulsar frå fleire tradisjonar. Ein slik løysingsstrategi vil dessutan bidra til fragmentering av forskingsfeltet. Metakriterium er problematiske dels av di dei eksisterer i mange (svært ulike) variantar, dels av di dei per definisjon er abstrakte og generelle og difor til liten hjelp i arbeidet med å evaluere konkrete prosjekt og tekstar. Sjekklister og andre forslag til meir handfaste evalueringskriterium er problematiske av di dei ikkje i tilstrekkeleg grad tar omsyn til at feltet faktisk er pluralistisk og at ulike artiklar difor treng ulike former for evaluering.

Problema er meir prinsipielle enn at dei lar seg løyse ved nye og betre forsøk på å artikulere lokale kriterium, metakriterium eller sjekklister. Kvalitativ forskning er slik sett truga av ein fare for fragmentert evalueringspraksis, vag evalueringspraksis og/eller evaluering på feil premiss. I artikkelen foreslår vi difor at regel- og kriteriebasert evaluering vert erstatta av ei agendabasert tilnærming, med den implikasjonen at evalueringspraksisen vert mindre ekspertstyrt og meir dialogisk i forma. Ei agendabasert tilnærming vil slik kunne gje rom for pluralisme på ein heilt annan måte enn ei kriteriebasert tilnærming. Ei legitimering av ei slik tilnærming til evaluering finn vi mellom anna i John Dewey's vitskapsfilosofi. Dewey argumenterte for at evalueringskriterium for forskning ikkje er universelt gyldige eller uavhengige av praksiskontekstar. Samstundes er slike standardar heller ikkje heilt lokale. Evalueringskriterium sprenger etter kvart rammene for det enkelte forskar-

samfunnet og får tyngde gjennom vedvarande bruk og kritikk i ulike samanhengar.

Vi argumenterer difor for at dialog og refleksivitet må vere berande prinsipp for evaluering av kvalitativ forskning. Som reiskap for løysingsorienterte evalueringssamtalar, lanserer vi då ein agenda som vi gjev namnet EPICURE. Ideen med ein agenda er mellom anna at bruken av han ikkje føreset semje om kunnskapsteoretiske spørsmål, berre om kva tema som (gjennom vedvarande bruk og kritikk) har nok tyngde til at dei er verdt å drøfte. EPICURE er eit akronym for eit forslag om sju slike tema: Engagement, Processing, Interpretation, Critique, Usefulness, Relevance og Ethics. Agendaen inneheld to dimensjonar, EPIC og CURE, der Critique (som høvesvis sjølvkritikk og sosial kritikk) er det overlappende punktet. Den første dimensjonen viser til forskinga si utfordring i høve til å utforme truverdige og engasjerande framstillingar av dei fenomena som vert studerte, medan den andre viser til forskinga sin relasjon til eit praksisfelt.

Artikkelen klårgjer desse omgrepa og premissa dei byggjer på og gjev også eksempel frå publisert forskning. Vidare vert sjølve agendaen diskutert, både i høve til korleis ulike tema kan ha ulik grad av relevans i høve til ulike prosjekt og tradisjonar og i høve til korleis bruken av han vil kunne føre til endringar i evalueringspraksis, ved at ein agenda fordrar eksplisitt refleksivitet ikkje berre hjå forfatarane men også hjå fagfellane.

### **Oppsummering og forslag til utprøving og vidare forskning**

Eg kan forsøkje å oppsummere argumentasjonen ved å bruke eit bilete: Om to store bilar køyrer i motsett retning på ein smal, svingete vestlandsveg, kan det hende at det rett og slett ikkje er mogeleg for dei å passere kvarandre. Prøver dei seg på dette likevel, køyrer dei seg gjerne fast mellom autovernet og fjellveggen. I så fall er det vegen meir enn dugleiken til sjåførane som er problemet (sjølv om vi nok kunne stille spørjeteikn ved dømekrafta deira). Tilsvarande er det kanskje slik innan kvalitativ forskning at kriteriebasert evaluering er å ta ein svært smal veg, der kommunikasjonen mellom fagfelle og forfatar står i fare for å køyre seg fast av di der ikkje er laga plass for ulike perspektiv. På same måte som problemet med smale vestlandsvegar gjerne vert omgått ved at den eine køyrer ut på ein møteplass og gjev den andre forrang, kan problemet med kriteriebasert evaluering omgåast ved å gje den eine sitt perspektiv forrang (då er det jo som oftast forfattaren som må gje seg). Dette er likevel ikkje ei fullgod løysing, då det legg dempar på kreativitet og nyutvikling. Ei agendabasert tilnærming til evaluering inneber at ein tar ein heilt annan veg, og då etter denne forfattaren si vurdering ein veg med meir rom for pluralisme og ulike perspektiv, altså ein veg som er

meir i samsvar med behova innan det vide feltet som kvalitativ forskning representerer.

No er det slik at vegar må byggjast, det er ikkje nok å teikne dei. Artikkelen som eg har referert til er strengt tatt berre ei vegteikning. Vegen må byggjast gjennom praksis, utprøving og ny forskning. Det kan kome til å koste litt, då der er mange tunge aktørar som har investert i kriteriebasert evaluering. Som døme kan eg nemne ein instans som Nasjonalt kunnskaps-senter for helsetjenesten og eit tidsskrift som *BMJ*, som begge har publisert sjekklister for vurdering av kvalitativ forskning (10,11). Det vil truleg i dag vere nokså vrient å kome særleg langt med ein kvalitativ artikkel innsendt til tunge og etablerte tidsskrift, dersom ein argumenterer for at ein ønskjer seg ei anna evaluering enn det sjekklista deira legg opp til. Ei agendabasert tilnærming til evaluering representerer såleis ei solid utfordring for fagfellevurderte tidsskrift, då ei dialogisk tilnærming fordrar ei viss maktomfordeling i evalueringsprosessen. Nett korleis ein meir agendabasert og dialogisk orientert evalueringspraksis kan utformast slik at han fungerer både for forfattarar, fagfellar, redaktørar og lesarar gjenstår å prøve ut, evaluere og forske på. I løpet av den korte tida som har gått sidan publisering av artikkelen referert til over, har eit par internasjonale musikkterapitidsskrift signalisert interesse for å gjere nett dette, og det er sjølv sagt ønskjeleg at fleire tidsskrift innan fleire ulike fagområde gjer det same.

Sett frå synsstad til ein forskar og forfattar er tidsskrifta sin praksis oftast lettare å handtere enn enkelte andre evalueringssituasjonar. Eit avslag frå eit tidsskrift opnar opp for bruk av eit anna tidsskrift, dersom det no er slik at ein opplever at evalueringa ikkje er rimeleg men byggjer på premiss som er framande for teksten. Ei negativ vurdering av eit doktorgradsprosjekt eller ein søknad på eksterne prosjektmidlar kan i større grad opplevast som ei stengd dør. Forsking på evalueringspraksisar innan desse områda kunne såleis vere svært viktig. Mange kvalitative forskarar klagar til dømes over at kvalitative søknader til forskingsråd og etiske komitear ikkje vert vurderte på rimeleg vis. Er det noko i dette? I tilfelle har kunnskapsutvikling basert på kvalitative metodar eit problem som aktørar innan eit fag som allmennmedisin vil kjenne igjen frå sin oppbyggingsperiode som akademisk fag. Steinar Westin har argumentert for at dei vitskapsteoretiske krava som vart stilte ved vurderinga av ei av dei første doktoravhandlingane i allmennmedisin på 1960-talet var urimelege i høve til forskingsfelt og kontekst og at denne avhandlinga difor vart avvist på feil premiss. Han går så langt at han kallar dette «sykehusmedisinens herjing med det allmennmedisinske feltet» (12).

Det er ikkje mi oppgåve her å ta stilling til om dette er ei presis beskriving, men eksempelet illustrerer at evaluering også er knytt til makt; nokre

stemmer får sleppe til, andre ikkje. Det er ikkje nødvendigvis snakk om urimeleg maktbruk. Det ligg i forkinga sin natur at nye funn og formuleringar skal prøvast kritisk, jamfør sosiologen Mertons (1910-2003) klassiske diskusjon av sentrale og vedvarande normer i vitskapen (13). Dersom vurderingar av kvalitativ forskning vert gjort på premiss som ikkje samsvarar med dei som undersøkingane byggjer på, kan det derimot vere snakk om urimeleg maktbruk. Slik vil til dømes sjekklister for vurdering av kvalitativ forskning kunne opplevast som grunnlag for urimeleg vurderingspraksis, dersom dei byggjer på premiss som ikkje passar, til dømes ved at dei i for stor grad er tatt over frå kvantitative forskningstradisjonar. Korleis slike sjekklister brukast i praksis er igjen eit spørsmål som kunne og burde ha vore undersøkt empirisk. Vert eventuelle slagsider eller manglar i sjølve sjekklis- tene kompensert for eller forsterka gjennom bruken av dei?

God forskning føreset grundige og gjerne utfordrande evalueringsspro- sar. På same måte som vi må vere merksame på risikoen for at evaluering på urimelege premiss kan avvise stemmer som burde ha fått sleppe til, må vi vere merksame på at fragmentert eller vag kriteriebasert evalueringsspraksis kan føre til at for mykje svak eller halvgod kvalitativ forskning vert publisert (14). Her kan det vere nokre innebygde utfordringar også for ei agendabasert tilnærming, dersom dialogens tendensar i retning konsensus får dominere på ein måte som ikkje gjev tilstrekkeleg rom for kritisk refleksjon. Utprøving av og forskning på ei agendabasert tilnærming til evaluering bør kunne ha eit kritisk blick på dette. Når eg i denne artikkelen har argumentert for det vi kan kalle ein romslegare evalueringsspraksis med meir rom for pluralisme, så er det altså ikkje det same som å argumentere for ein slakkare evalueringsspraksis. Forslaget om ei agendabasert tilnærming til evaluering av kva- litativ forskning må snarare forståast som ei oppmoding til «kontinuerleg vegbygging» for betre evaluering av betre forskning.

## Litteratur

1. Malterud K. Qualitative research: standards, challenges, and guidelines. *Lancet* 2001; 358: 483-8.
2. Malterud K. The art and science of clinical knowledge: evidence beyond measures and numbers. *Lancet* 2001; 358: 397-400.
3. Stige B, Malterud K, Midtgarden T. Toward an agenda for evaluation of qualitative research. *Qualitative Health Research* 2009; 19: 1504-16.
4. Ruud E. Om forskning i musikkterapien: metakritikk, verdier og kvalitativ metode. *Nordic Journal of Music Therapy* 1992; 1: 21-6.
5. Stige B. Do we need general criteria for the evaluation of qualitative research articles, and if we do, how could such criteria be formulated? *Nordic Journal of Music Therapy* 2002; 11: 65-71.

6. Wittgenstein L. *Philosophical investigations*. Oxford: Blackwell, 1953/1967.
7. Alvesson M, Sköldbberg K. *Reflexive methodology: New vistas for qualitative research*. London: Sage Publications, 2000.
8. Finlay L., Gough, B. (red). *Reflexivity. A practical guide for researchers in health and social sciences*. Oxford, UK: Blackwell Publishing, 2003.
9. Aigen K. Philosophical inquiry. I: Wheeler B red. *Music therapy research* (2.utgave). Gilsum, NH: Barcelona Publishers, 2005: 526-39.
10. Nasjonalt kunnskapssenter for helsetjenesten. Sjekklister for vurdering av forskningsartikler [online], publisert 14.05.2008. Tilgjengeleg frå: <http://www.kunnskapssenteret.no/Verkr%C3%B8y/2031.cms>
11. *BMJ*. Resources for authors, special methods [online]. Lasta ned 20.08.2007, frå: <http://resources.bmj.com/bmj/authors/special-methods>
12. Westin S. Allmennmedisin som akademisk fag. *Michael* 2009; 6: 65.
13. Merton RK. *The Sociology of science. Theoretical and empirical investigations*. Chicago: The University of Chicago Press, 1942/1973.
14. Pope C & Mays N. Critical reflections on the rise of qualitative research. *BMJ* 2009; 339: b3425 (September), 737-9.

*Brynjulf Stige*  
*Griegakademiet – Institutt for musikk*  
*Universitetet i Bergen*  
*Brynjulf.Stige@grieg.uib.no*