

Hva om medisinstudenter lager oppgaver til egen eksamen?

I undervisningen ved Seksjon for allmennmedisin i Bergen ønsker vi å flytte vekten fra kontroll til læring. Som et ledd i dette arbeidet tilbød vi høsten 2000 avgangskullet å lage oppgaver til sin egen skriftlige eksamen i allmennmedisin. Vi lovet studentene å velge ut én oppgave for hvert tredje forslag vi fikk inn, eventuelt noe bearbejdet. Ved fristens utløp forelå det tre forslag, og ett av disse ble brukt i eksamenssettet. Målsettingen var å evaluere konsekvensene av å la medisinstudenter skrive oppgaver til egen eksamen.

Tiltaket ble evaluert ved eksamenskarakter, studentenes skåre på oppgavens relevans og om prosessen hadde fått dem til å endre lesevaner og fritekstevaluering av de samme to forholdene.

57 av 64 (89 %) eksamenskandidater leverte inn evalueringsskjema. Alle visste før eksamen at én av eksamensoppgavene ville være basert på studentenes forslag, men bare 34 (60 %) angav at dette påvirket lesingen. Alle mente at studentoppgaven var minst like relevant for allmennmedisin som de andre oppgavene. Gjennomsnittlig karakter var 9,9 (variasjonsbredde 5–12) på studentoppgaven, mot 9,5 (5–11) på skriftlig eksamen som helhet, middeldifferanse 0,32, 95 % konfidensintervall 0,08–0,64. Fritekstevalueringen viste at studentene leste ekstra nøye på de temaene oppgaveforslagene omhandlet. Det gav trygghet i alle fall å kjenne én av oppgavene til eksamen.

Studentene angav at studentoppgaver til eksamen gjør eksamenssituasjonen tryggere, mens relevansen for allmennmedisin opprettholdes. Eksamens kontrollfunksjon synes å være bevart.

Eksamen ved medisinstudiet har flere funksjoner. Den skal fungere som kontroll på at studenten har tilegnet seg det minstemål av kunnskaper, ferdigheter og holdninger som ansees som nødvendig før turnustjeneste. I så måte burde man kunne kreve at eksamen har både god validitet og god reliabilitet, det vil si gjenspeiler studentens prestasjoner både riktig og nøyaktig. Tilgjengelig dokumentasjon sår sterk tvil om det er slik (1, 2).

Anders Bærheim

anders.barheim@isf.uib.no

Eivind Meland

eivind.meland@isf.uib.no

Seksjon for allmennmedisin

Institutt for samfunnsmedisinske fag

Universitetet i Bergen

Ulriksdal 8C

5009 Bergen

Bærheim A, Meland E.

What happens when medical students set their own exam papers?

Tidsskr Nor Lægeforen 2001; 121: 2971–3.

Background. In the section for general practice at the University of Bergen, Norway, we want to emphasize learning more than control in our work with students. As a step in this direction we invited students to come up with proposals for exam papers, papers that usually include six to nine multiple steps for clinical reasoning. We guaranteed that one out of three proposals would be included as a paper in the written examination, possibly slightly modified. This article is an evaluation of the consequences of letting medical students set some of their own exam papers.

Material and methods. The process was evaluated using 1) grades given, 2) students' assessment of whether this mode of setting papers influenced their exam preparations, and 3) students' free-text comments on the process.

Results. 57 out of 64 students (89 %) took part in the evaluation. All knew that their fellow students had set one of the exam papers, but only 34 (60 %) reported that this knowledge had changed the way they prepared for the exam. The mean grade was 9.9 (range 5–12, on a scale from 1 to 12, 6 being the lowest pass grade) for the paper set by students, and 9.5 (range 5–11) for all papers combined. Mean difference in score was 0.32 (95 % confidence interval 0.08–0.64). Students' free-text comments showed that they specifically prepared for the three known paper topics. They drew comfort from knowing at least one of the papers set, and the student-set papers were found relevant for general practice.

Interpretation. Letting medical students set one of the exam papers makes them feel more confident. Student-set papers were seen as relevant for clinical practice. The control function of the exam seemed to have been preserved.

Eksamens form og innhold styrer studentenes studieatferd mer enn noe annet element i studieplanen (2). Studentenes forventning til en tradisjonell skriftlig eksamen kan motivere dem til overflatelæring og reproduksjon av kunnskap, og i mindre grad inspirere til refleksjon, dybdeforståelse eller ferdighetstrening (3). Legendariske sir Wil-

liam Osler skrev allerede i 1913 om eksamen: «Nothing is more fatal to a true intellectual training than a constant preoccupation with its practical results» (4).

Medisinstudiet i Bergen har i alt 26 eksamener. Det legges ned en betydelig mengde arbeid i gjennomføringen av disse eksamenene, også fra studentenes side. Eksamen er forbundet med stress og usikkerhet hos studentene (2, 5), og spesielt er de åtte eksamenene som er hopet opp ved avslutningen av studiet i Bergen en stor belastning.

Det kunne være ønskelig å minske denne belastningen på studentene. Samtidig er det innen høyere utdanning, og ikke minst innen medisinstudiet, behov for å flytte oppmerksomhet og ressursbruk fra kontroll til læring. Strategisk plan 2000–05 for Universitetet i Bergen fastslår at tid og ressurser som går med til eksamen og sertifisering må reduseres, og at studentene i stedet bør forpliktes til læring blant annet gjennom mer akademisk skriving (6).

Som et element i dette arbeidet innførte vi høsten 2000 en ordning der studentene fikk tilbud om å lage oppgaver til sin egen skriftlige eksamen i allmennmedisin. Så langt vi har brakt i erfaring, har dette ikke vært gjort før. Riktignok rapporterer Sirclar & Tandon fra et prosjekt der studenter var med på å lage flersvarsoppgaver til en oppgavebank (7), men disse studentene fikk ikke løfte om at de ville få oppgavene de produserte til sin egen eksamen.

Vi tilbød ordningen til avgangsstudentene høsten 2000. De leverte tre oppgaveforslag, og ett av disse ble inkludert som oppgave i eksamenssettet. Siden ordningen er et pedagogisk nybrottsarbeid, ønsket vi å evaluere i hvilken grad dette påvirket studentenes læringsatferd før eksamen, i hvilken grad de oppfattet studentoppgaven som relevant for allmennmedisin, samt hvordan de ville skåre ved eksamen på selvprodusert oppgave i forhold til resten av eksamenssettet.

Materiale og metode

Ved medisinstudiet i Bergen undervises det i allmennmedisin vesentlig gjennom en åtte ukers termin i siste studiehalvår, seks uker før avsluttende eksamen. Avsluttende eksamen består av en femtimers skriftlig eksamen i forebyggende medisin og sosialmedisin, samt fem timer skriftlig og 1–2 timer muntlig i hver av de tre kliniske hovedfagene kirurgi, indremedisin og allmennmedisin.

Medisinstudiet i Bergen har hatt eksamen i allmennmedisin siden 1994. Skriftlige

Den studentproduserte oppgaven til eksamen i allmenntilleggsfag i medisin høsten 2000 (8). Tillegg føyd til av eksamenskommissjonen er fremhevet i kursiv. Studentenes opprinnelige forslag var: a. Hva vil du legge vekt på i den videre anamnesen?

Lise Lyshø, 23 år, kommer til en øyeblikkelig hjelp-time mandag formiddag. Hun forteller at det har svidd når hun tisser de siste tre dagene. Hun har også ubehag i underlivet og føler seg slapp. Hun hadde blærekatarr for et år siden, og kjenner igjen symptomene. Under konsultasjonen forteller hun at hun fikk utført en provosert abort for 10 dager siden. Hun har ingen barn fra før.

a. *Du undersøker urinen med stiks og finner at den er positiv på leukocytter. Fabrikanten angir at leukocytffeltet på deres urinstrimmel har en sensitivitet på 0,85 og en spesifisitet på 0,80. Du vet at typiske cystittsymptomer har en pretestsannsynlighet for cystitt på 0,90 når kvinnen kjenner igjen symptomene fra tidligere cystitter. Hva blir posttestsannsynligheten for cystitt, gitt positivt utslag på leukocytffeltet? Svar ved en 2×2 -tabell eller Fagans nomogram (vedlagt).*

b. *Er det opplysninger i den aktuelle sykehistorien som vil påvirke posttestsannsynligheten for cystitt som du har beregnet? Hvis ja, hvordan?*

c. *Hva vil du legge vekt på i den kliniske undersøkelsen av Lise Lyshø, og hvilke supplerende undersøkelser vil du utføre?*

Pasienten ønsker prevensjon. Hun har tidligere brukt p-piller, men har lest i avisen at man kan få blodpropp av å bruke dem. Hun sluttet med det og ble gravid. Hun lurer på om du vil anbefale minipiller eller spiral.

d. *Hvordan vil du informere Lise om risikofaktorer for blodpropp ved bruk av p-piller, og hvilke opplysninger vil du ta med?*

e. *Sett opp skjematisk forskjeller mellom p-piller og minipiller. Hva vil du anbefale henne og hvorfor?*

f. *Hva slags typer spiral finnes det? Ville du anbefalt henne dette? Begrunn svaret.*

eksamensoppgaver fra tidligere år ligger i studentportalen på hjemmesiden vår (8). Mens muntlig eksamen i allmenntilleggsfag i medisin er en konsultasjon (9), setter vi ved skriftlig eksamen søkelys på klinisk håndtering av allmenntilleggsfag i medisin tilstander. Oppgavene krever at studentene har kunnskap om klinisk epidemiologi og kan resonere ut fra både sannsynlighetsbaserte og etiske overveielser. Oppgavesettene ved skriftlig eksamen i allmenntilleggsfag i medisin har de siste årene hatt 6–14 oppgaver med flere ledd.

Høsten 2000 oppfordret vi ved terminstart studentene til å skrive sine egne eksamensoppgaver. Vi lovet dem at hvis de laget tre eller flere forslag til eksamensoppgaver, ville hvert tredje forslag bli gitt som oppgave ved deres egen skriftlige eksamen i allmenntilleggsfag i medisin samme høst.

Studentene ble spesielt gjort oppmerksom på at forslagene til oppgave kunne baseres på deres egne kliniske erfaringer, for eksempel konsultasjonen som de tok opp på video til bruk i konsultasjonskurset vårt (10). Studentene leverte inn tre forslag. Alle tre forslagene var rimelig gjennomarbeidet, og vi fant at de omhandlet sentrale allmenntilleggsfag i medisin tema. Temaene var, sterkt resymert: – Stine, fem år, som våknet med gjenklistret høyre øye. Tilleggsopplysninger ble gitt, og det ble spurt etter mulige diagnoser, hvordan pasienten kunne behandles, hvilken informasjon som kunne gis, og om hvilke forebyggende tiltak som kunne være aktuelle.

– En 57 år gammel kvinne med influensaliknende sykdom, etterfulgt etter en uke av et uklart kløende utslett. Spørsmål gikk på mulige diagnoser, behov for tilleggsopplysninger, begrunnet diagnose, behandling og prognose.

– Oppgaven som ble valgt (ramme).

Etter endt eksamen ble studentene bedt om å evaluere ordningen ved hjelp av et anonymt spørreskjema. Skjemaet ble av praktiske årsaker levert ut mens studentene ventet på muntlig eksamen i allmenntilleggsfag i medisin. Eksaminator og sensor så ikke skjemaene før etter avsluttet muntlig eksamen. Studentene ble ved kategoriske spørsmål spurt om de visste at slike eksamensoppgaver ble laget (skåre ja/nei), om det påvirket lesingen til eksamen (skåre ja/nei), om hvordan studentoppgaven fungerte i eksamenssituasjonen i forhold til de andre oppgavene (skåre fra 1 = mye dårligere til 5 = mye bedre) og hvor relevant den var for allmenntilleggsfag i medisin (skåre fra 1 = mindre til 3 = mer). Vi bad dem også angi i fritekst hvordan tiltaket hadde endret lesingen deres samt hvordan de bedømte oppgavens relevans for allmenntilleggsfag i medisin praksis.

I den kvantitative delen av bearbeidningen brukte vi resultatet av eksamenskommissjonens arbeid som utfall. Eksaminator og sensor satte konkret karakter på hver oppgave ved retting, og karakteren på studentens oppgave kunne sammenliknes med total karakteren på eksamenssettet. Verken eksami-

nator eller sensor var blindet for hvilken oppgave som kom fra studentene.

Databehandling

Forskjeller i eksamenskarakter er angitt ved snitt med variasjonsbredde, og middel differanse med 95 % konfidensintervall. Fritekst ble analysert fenomenologisk ved en variant av Giorgis metode etter modifikasjon ved Malterud (11).

Resultater

I alt 57 av de 64 studentene som gikk opp til eksamen høsten 2000 leverte inn evalueringsskjema. Alle 57 visste at én av eksamensoppgavene ville være basert på studentenes forslag, og 34 studenter angav at dette påvirket lesingen. I alt 45 studenter mente at studentoppgaven fungerte like bra som eksamensoppgave som de andre oppgavene, tre studenter mente at den fungerte dårligere og seks at den fungerte bedre. 48 studenter fant at deres oppgave var like relevant for allmenntilleggsfag i medisin praksis som de andre oppgavene, og fem studenter mente at den var mer relevant.

Studentene skåret gjennomsnittlig 9,9 (variasjonsbredde 5–12) på studentoppgaven, mot 9,5 (variasjonsbredde 5–11) på skriftlig eksamen som helhet. Forskjellen (middel differanse 0,32, 95 % konfidensintervall 0,08–0,64) var statistisk signifikant.

I alt 49 studenter leverte evaluering i fritekst. I den kvalitative analysen av denne teksten identifiserte vi følgende tre kategorier.

Endret leseatferd

De fleste studentene oppgav å ha lest alle tre oppgavene som medstudentene deres hadde sendt inn, og de hadde arbeidet med å løse dem både individuelt og i kollokviegrupper. En del av studentene leste også mer på nærliggende områder innen temaene hud, oftalmologi og gynekologi, siden de visste at den oppgaven som ble valgt, kunne omformes noe. Dette gav trygghet: «(trygt) å vite at i alle fall én av oppgavene var innen et kjent område.»

Flere andre var trygge nok til ikke å lese så mye ekstra på de emnene eksamensforslagene representerte. En skrev: «Vi har vel etter hvert forstått hvordan eksamensoppgaver utformes.»

Andre oppgav at de ville lest på de samme temaene uansett, siden de var sentrale, eller at de allerede hadde lest dem da de så oppslaget med oppgavene på studenthuset. En student rapporterte at flere ble stresset dagen før eksamen av å vite hva som ville bli gitt.

Relevante oppgaver

De fleste studentene mente at alle tre oppgavene fra studentene var svært relevante, både den som ble gitt til eksamen og de to andre. De beskrev tilstandene som hyppige i allmenntilleggsfag i medisin praksis, og mange hadde opplevd liknende konsultasjoner selv. Flere satte

eksplisitt pris på at oppgaven baserte seg på en genuin opplevelse fra en medstudent: «Historien var sann, så den er veldig relevant.»

«Gav oss mulighet til å vise hva vi kunne»

Studentene mente oppgaven var slik at de fikk vist hva de kunne. De likte vekslingen mellom konkrete kunnskapsspørsmål og spørsmål som krevde resonnerende svar. Spørsmålene krevde resonnementer med utgangspunkt i kunnskap om klinisk epidemiologi så vel som vekt på pasientens forventninger, forståelse og følelser. Samtidig var mange studenter fornøyd med at oppgaven også inneholdt spørsmål som kunne besvares som kunnskapsreproduksjon. En skrev: «Varierte problemstillinger fra det rent medisinske til kunnskapsbasert medisin og psykosomatikk.»

Diskusjon

Det er ofte vanskelig å vurdere effekten av pedagogiske tiltak, ikke minst fordi få mål er valide for hvordan studentene forholder seg til tiltaket i praksis og hvilket læringsutbytte det gir. Karakterer til skriftlig eksamen har både reliabilitets- og validitetsproblemer i forhold til den individuelle kandidaten (2), og er et dårlig mål for endringer i studentenes læringsprosess.

Det var praktisk umulig å blinde eksaminatorer og sensorer med hensyn til hvilken oppgave studentene hadde laget. Eksaminatorene satte selv sammen eksamenssettet og utarbeidet rettegilde sammen med sensorene. Dermed kan vi ikke utelukke at studentoppgaven ble rettet med en annen holdning enn de andre oppgavene. På den annen side inneholdt studentoppgaven mange konkrete spørsmål. Dette muliggjorde en spesifikk rettegilde med mindre rom for skjønn.

Vi la i det ene fritekstspørsmålet vårt vekt på endrede lesevaner. Litt over halvparten av studentene angav at de leste mer på de emnene oppgaveforslagene fra studentene omhandlet. Viktigere var det nok at de leste annerledes. Mange la en strategi for løsning av oppgaveforslagene, og øvde seg konkret i måter å løse dem på. En slik prosess vil kunne bidra til økte ferdigheter i bredt klinisk resonnement så lenge oppgavene vektlegger det.

Tidligere eksamensoppgaver med slikt innhold er utgangspunkt for studentenes selvdrevne eksamenskollokvier, og gir dermed rom for trening i klinisk resonnement. Som én av studentene skrev – de vet hvordan eksamensoppgavene i allmenntid medisin utformes, og enhver eksamensform sosialiserer studentene til å øve inn de svarteknikkene som eksamen krever (2, 3). Våre resultater indikerer at studentenes eksamensforberedende arbeidsvaner er labile, og at selv små endringer i eksamensform kan medføre endringer i disse.

Tiltaket vårt synes videre å kunne være et

skritt i retning av å ufarliggjøre eksamen, idet flere studenter angav at de følte seg tryggere når de i alle fall kjente én av oppgavene. Vi kan ikke uten videre hevde at vår målsetting om å flytte oppmerksomheten fra kontroll til læring er oppnådd. De atferdsendringene studentene beskriver, ble tross alt induert av vår endrede eksamensform. Vi vet ikke om den intensiverte læringen innen de tre gitte emnene skjedde på bekostning av andre kliniske temaer.

Den viktigste effekten av tiltaket vårt blir å finne i det arbeidet som ble utført av de studentene som leverte inn de tre forslagene. Det innebærer et viktig refleksjonsarbeid når man skriver sitt fag (7, 12, 13), og studentene som leverte inn oppgavene, må antas å ha lært en god del gjennom aktiv klinisk refleksjon mens de utarbeidet forslagene.

Vi håper derfor å få flere studenter på kommende kull med i prosessen. Det ville vært ideelt om alle de 60 studentene i et kull gikk sammen to og to eller tre og tre om å skrive oppgaver. Da ville vi fått inn 20–30 forslag, og samtlige 6–10 oppgavene i vår skriftlige eksamen kunne være hentet fra studentenes forslag. Studentene ville ha brukt mye tid til å reflektere klinisk under utformingen av forslagene, de ville lest på hele bredden av faget, som 30 oppgaver uvegerlig vil omfatte, samtidig som eksamenskontrollfunksjon ville vært rimelig intakt.

Enkelte vil kanskje hevde at eksamen ville bli for enkel og dårlig egnet som kontroll av studentenes kunnskaper om flere av oppgavene var laget av studentene selv. Vårt resultat viser at studentene skåret en tredels karakter høyere på deres egen oppgave enn på eksamen som helhet. Dette er neppe av praktisk betydning. Imidlertid kan det gjenspeile at studentene tross alt har arbeidet mer med problemstillingene. Leverer studentene flere oppgaveforslag, vil flere emner innen faget kunne få en tilsvarende grundigere behandling. Eksamens egnethet som kontroll og godkjenningssredskap vil sannsynligvis bli opprettholdt, selv om eksamenskarakteren kan bli noen tideler høyere. Det er imidlertid viktig at studentene er klar over at eksaminatorene kan reformulere oppgaver og supplere med egne spørsmål. For å bevare eksamens kontrollfunksjon og forebygge overflatelæring kan det være nødvendig å presisere for studentene at deres egne løsningsforslag kan være irrelevante.

Studentenes innsats i formuleringen av oppgaver vil gi dem økt trening i klinisk refleksjon. Skrivning av eksamensoppgaver vil være et sterkere incitament i så måte enn vanlig kollokvering med oppgaveløsning. Vi antar at tiltaket innebærer at ressurser blir overført fra kontroll til læring, selv om denne studien ikke kan dokumentere dette. En slik ressursoverføring fordrer sannsynligvis at storparten av studentene engasjerer seg i oppgaveformulering.

Vi vil fortsette forsøksordningen med å benytte studentproduserte kliniske oppgaver

til skriftlig eksamen i allmenntid medisin. Resultatene av evalueringen viser at eksamenssituasjonen blir tryggere for studentene. De vurderte sine egne oppgaver som like relevante for faget, og forskjellene i karakter har ikke et nivå som underminerer eksamens kontrollfunksjon.

Litteratur

1. Brodal P. Eksamen – myter og realiteter. Tidsskr Nor Lægeforen 1997; 117: 489–90.
2. van der Vleuten C. Validity of final examinations in undergraduate medical training. BMJ 2000; 321: 1217–9.
3. Kvale S. Eksamen som konstruksjon av kunnskap. Uniped 2000; 22: 6–22.
4. Osler W. An introductory address on examinations, examiners, and examinees. Lancet 1913; 2: 1047–50.
5. Godfrey RC. Undergraduate examinations – a continuing tyranny. Lancet 1995; 245: 765–7.
6. Strategisk plan 2000–2005 for Universitetet i Bergen. Bergen: Universitetet i Bergen, 1999: 22.
7. Sirclar SS, Tandon OP. Involving students in question writing, a unique feedback with fringe benefits. Am J Physiol 1999; 277: S84–91.
8. <http://www.uib.no/isf/student/> (28.2.2001).
9. Bærheim A, Malterud K. Klinisk eksamen i allmenntid medisin – en konsultasjon. Erfaringer fra førstegangseksamen i Bergen. Tidsskr Nor Lægeforen 1995; 115: 1368–70.
10. Schei E, Bærheim A, Meland E. Klinisk kommunikasjon – et strukturert undervisningsopplegg. Tidsskr Nor Lægeforen 2000; 120: 2258–61.
11. Malterud K. Kvalitative metoder i medisinsk forskning. Oslo: Tano, 1996: 92–102.
12. Deloney LA, Carey MJ, Beeman HG. Using electronic journal writing to foster reflection and provide feedback in an introduction to clinical medicine. Acad Med 1998; 73: 574–5.
13. Buchanan H, Moore R, van Niekerk L. The fieldwork case study: writing for clinical reasoning. Am J Occup Ther 1998; 52: 291–5.

