

HPV-testing – svak begrunnelse for kontroversiell beslutning

Det har vært en langvarig konflikt om bruk av en norskutviklet mRNA-test i sekundærscreeningen for livmorhalskreft. Testen er i praksis utelukket fra det norske markedet etter at departementet vedtok at den fra 1.1. 2012 ikke gir rett til refusjon. I det følgende vil jeg drøfte noen sider av konflikten og argumentene for at forvaltningen aldri har gitt en holdbar begrunnelse for sin beslutning.

 Publisert først på nett 30.5. 2012

Ivar Sønbo Kristiansen

i.s.kristiansen@medisin.uio.no

Avdeling for helseledelse og helseøkonomi
Universitetet i Oslo

I 2005 anbefalte screeningprogrammet for livmorhalskreft (1) bruk av HPV-test i etterundersøkelse av kvinner som hadde usikre celleforandringer i den primære celleprøven. De ulike sykehuslaboratoriene tok i bruk enten én av flere DNA-tester eller en norskprodusert mRNA-test. I 2010 foreslo Helsedirektoratet å fjerne refusjon for alle HPV-tester (2), men helseministeren tok ikke forslaget til følge. Våren 2011 foreslo Helsedirektoratet nye krav til HPV-tester (3), krav som i praksis ville fjerne refusjon for mRNA-testen. Helse- og omsorgsdepartementet sendte forslaget på høring, uten at det skjedde konkrete endringer. I desember 2011 kom en rapport fra Kreftregisteret om erfaringer med HPV-testing i sekundærscreening (såkalt triage) (4). Omtrent samtidig foreslo Helsedirektoratet igjen nye krav til HPV-tester (5). Departementet sluttet seg til forslaget, og mRNA-testen mistet refusjonen. Departementets håndtering av denne saken har gjentatte ganger vært tatt opp i Stortinget, senest i 2012, da helseministeren begrunnet sitt vedtak med Kreftregisterets rapport (6). Produsenten av mRNA-testen har reist rettsak mot departementet.

I 2009 ble jeg spurt om å delta i en arbeidsgruppe som vurderer innføring av HPV-testing i primærscreening. Jeg har derfor fulgt saken om valg av HPV-tester på forholdsvis nært hold.

Sensitivitet og spesifisitet

Infeksjon med HPV-virus kan forårsake cervikal intraepitelial neoplasia (CIN), som i noen tilfeller utvikler seg til invasiv kreft. CIN-tilfeller kan spores med celleprøve (cytologi) og diagnostiseres med biopsi (histologi). De fleste infeksjoner og celleforandringer er forbigående. I alders-

gruppen 20–30 år har 25 % av kvinner en pågående HPV-infeksjon, men andelen synker med økende alder (7). Antall krefttilfeller er knapt 300 per år (8), men ut fra mine beregninger hadde tallet trolig vært 500–750 uten screening (9).

Det tas årlig ca. 400 000 celleprøver i screeningprogrammet (10). Mer enn 30 000 kvinner får tatt kontrollprøver (celleprøve og/eller HPV-test), vel 5 000 får utført kolposkopi med biopsi hos gynekolog (10). Dette kan skape bekymring eller angst (11). Ca. 3 000 kvinner får utført konisering, hvorav vel halvparten er under 35 år (10). Konisering medfører økt risiko for senabort og for tidlig fødsel i senere svangerskap (12, 13). Programmets hjemmeside eller invitasjonsbrev for screening inneholder ikke informasjon om risikoen for falskt positive prøver eller om bivirkninger av konisering (14).

Det har vært benyttet flere ulike HPV-tester i sekundærscreening hos kvinner med usikre eller lavgradige celleforandringer. Den norske mRNA-testen har lavere sensitivitet for å oppdage CIN-tilfeller, men høyere spesifisitet, og gir derfor færre falskt positive prøvesvar (15). DNA-tester er mer sensitive, men mindre spesifikke (15).

Argumenter

Motstanderne av mRNA-testen henviser til at den har lavere sensitivitet for cervikal intraepitelial neoplasia (5). Tilhengerne hevder at mRNA-testen forebygger flere krefttilfeller fordi den oppdager onkogen aktivitet, ikke bare tilstedeværelse av HPV-virus (16). De hevder også at RNA-testen i praksis forebygger flere krefttilfeller fordi antall positive prøvesvar er lavere, slik at legene lettere kan følge opp pasientene.

Helsedirektoratet har stilt som krav til HPV-tester at de skal være validert i randomiserte kliniske studier (5). I Medline-søk fant jeg imidlertid bare én randomisert studie av HPV-testing i sekundærscreening (17). I denne studien, som er mindre relevant for norsk praksis, var DNA-testing ikke sikkert bedre enn cytologisk undersø-

kelse. Det er publisert en rekke randomiserte studier av DNA-tester i primærscreening, men disse er mindre relevante for sekundærscreening.

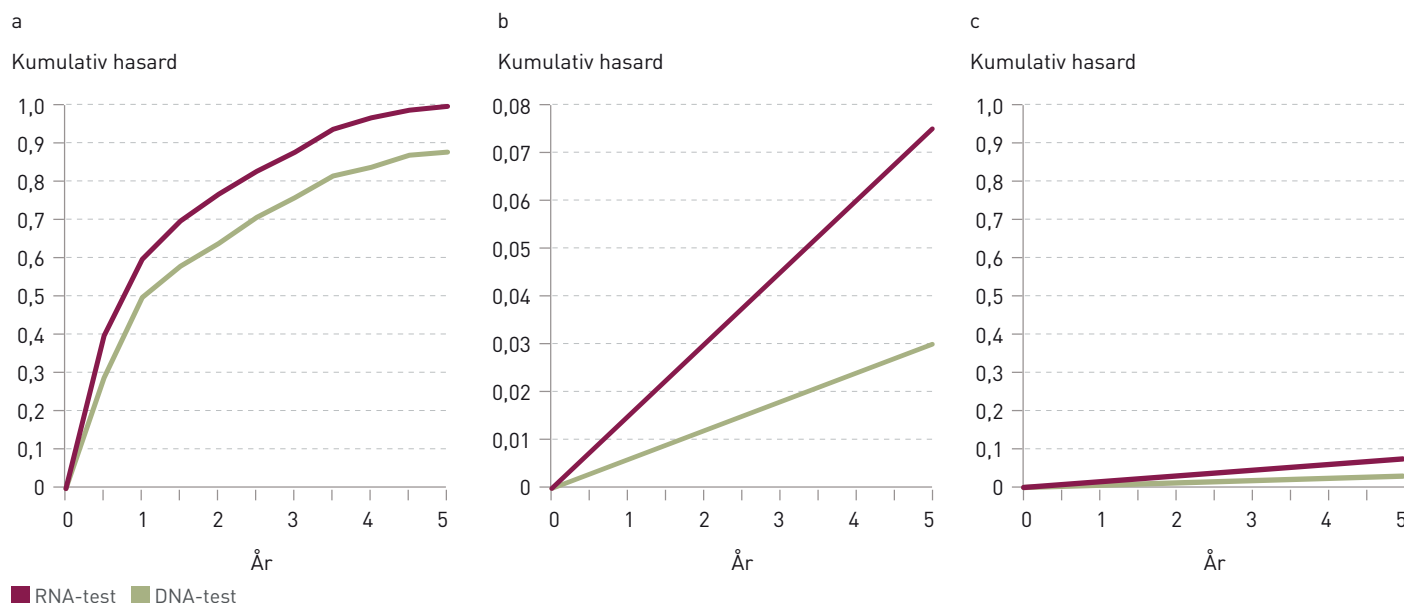
Valg av screeningstrategi krever at man tar hensyn til så vel falskt negative prøver (risiko for å overse kreft under utvikling) som falskt positive (risiko for etterkontroll av eller konisering for lesjoner som ikke vil utvikle seg til kreft). Man kan ikke a priori anse en test som bedre fordi sensitiviteten er høyere uten å ta hensyn til spesifisiteten. Forvaltningen synes imidlertid ikke å tillegge falskt positive prøver og koniseringsbivirkninger noen vekt.

Ett av Helsedirektoratets krav til tester, og dermed til refusjon, er at de skal oppdage minst 12 serotyper av HPV (5). mRNA-testen oppdager bare fem, men det er de antatt viktigste serotyper. Inkludering av flere betyr at en test blir mer sensitiv, men samtidig mindre spesifikk (18). Kravet om 12 serotyper er derfor en variant av kravet om høy sensitivitet, uten å avveie dette mot spesifisiteten.

Pseudoargumenter

Jeg har flere ganger etterlyst en konsistent begrunnelse for de krav forvaltningen har satt til HPV-tester (3, 5). Nedenfor angis noen av begrunnelsene jeg har mottatt i ulike e-poster. Forvaltningen har argumentert med at det er «bred faglig enighet» om at DNA-tester er best – selv om denne saken tvert imot er preget av uenighet. «Internasjonal konsensus» har vært et annet argument, uten at det har vært angitt hvordan man definerer eller måler et slikt begrep. Når man i blant annet Danmark rent faktisk bruker RNA-tester, kan man knapt snakke om konsensus.

En artikkel av Meijer og medarbeidere har vært brukt som argument for å utelukke mRNA-testen (19). Denne omtaler imidlertid valg av tester i primærscreening hos kvinner over 30 år, mens den norske debatten gjelder sekundærscreening. For øvrig er Meijer medlem av «advisory board» for Qiagen, produsent av én av DNA-testene.



Figur 1 Stilisert og forenklet fremstilling (ikke reelle tall) av figur 14 i Kreftregisterets rapport (4). a) Forekomst (kumulativ hasard) av CIN2+ blant kvinner i triage (sekundærscreening) for lavgradige celleforandringer påvist i primærscreening. b) Det samme i tiden etter triage – kan tolkes som uttrykk for CIN2+-tilfeller som var oversett i triageperioden. c) Viser det samme som b, bortsett fra at y-aksen har samme skala som i a. Visuelt kan b gi et inntrykk av at det er stor forskjell i hvor mye CIN2+ de to testene overser i triageperioden. Et annet visuelt inntrykk gir c. Også når panelene har samme skala på y-aksen, lar de seg vanskelig tolke fordi det ikke er angitt antall kvinner som inngår i grafene eller hvor stort bortfallet er over tid. Figur 14 i Kreftregisterets rapport og alle de tilsvarende figurene mangler signifikanstester for forskjeller mellom gruppene

Ett av argumentene mot mRNA-testen har vært at produsenten anbefaler ny celleprøve også etter negativ HPV-test hos kvinner med gjentatte celleforandringer. Dette er i strid med Kreftregisterets retningslinjer. I praksis blir imidlertid pasientene nesten like ofte fulgt opp med ny celleprøve etter negativ DNA-test (64 % mot 50 %) (20). Manglende etterlevelse av retningslinjene er utbredt i hele screeningprogrammet, ikke bare ved bruk av RNA-test (4).

Surrogatendepunkter

Veien til forebygging av livmorhalskreft er brolagt med usikkerhet. Vurderingen av celleprøver varierer med den som mikroskopiserer (21, 22). Patologene kan både over- og underdiagnostisere CIN-tilfeller i biopsier (23). Vurdering av kolposkopiske funn varierer hos gynekologene (24), og risikoen for ikke å treffe den aktuelle lesjonen kan være betydelig (25). Selv om den histologiske diagnosen alvorlig cervical intraepitelial neoplasi (såkalt CIN2+) er korrekt, betyr det ikke at lesjonen vil utvikle seg til kreft. Tvert imot utvikler bare mindretallet av CIN2+-tilfellene seg til kreft (26, 27). Hele screeningprogrammet er således basert på surrogatendepunkter, sett i forhold til det vi ønsker å forebygge: Invasiv kreft. Derfor bør man være varsom med å påstå at én test er bedre enn en annen.

Seleksjonsskjevhet

Rapporten fra Kreftregisteret (4) har en rekke svakheter. Bedømt etter fordelingen av cytologiske diagnoser (rapportens tabell 18) var pasientgruppene med mRNA-test

nokså forskjellige fra dem med DNA-tester. Dette gir mistanke om seleksjonsskjevhet og at man står i fare for å «sammenlikne epler og pærer». I rapporten står det derimot at «vi er ikke kjent med at det foreligger slike forskjeller», uten å vise dokumentasjon på at det ikke var forskjell mellom gruppene når det gjelder inntekt, sosial klasse, seksualvaner eller andre variabler som kan variere med kreftforekomsten.

Det kan tenkes at den nevnte skjevfordelingen skyldes at ulike laboratorier bruker ulike kriterier for å sette cytologidiagnoser. Det vil i tilfelle reise nye spørsmål om verdien av rapporten.

Hvilken test oppdager flest CIN-tilfeller?

Forfatterne av rapporten har analysert sine data både med krysstabeller og overlevelsesanalyser. Overlevelseskurvene (figurene 17–23) tyder på at det ble oppdaget flere CIN2+-tilfeller med mRNA-test enn med DNA-tester de første tre årene etter testing. Det motsatte var muligens tilfellet i noen grupper etter 3–5 års oppfølging, men her var tallene små og usikre.

Et interessant fenomen er bruken av grafer med ulike skala på y-aksene (fig 1) (4). Teknikken er egnet til å overdrive forekomsten av «oversett» CIN2+. Kreftregisteret ga ikke tillatelse til å gjengi originalfiguren her – selv om den allerede ligger på nettet.

I vel designede studier tar man biopsi også fra et tilfeldig utvalg av kvinner med negative tester for å korrigere for verifikasjonsskjevhet. Slik korrigerings kan ikke foretas med registerdata, men dette maner til varsomhet med datafortolkningen.

Kreftregisterets rapport inneholder praktisk talt ingen tester for om påviste forskjeller er statistisk signifikante. Noen forskjeller er åpenbart signifikante pga. de store tallene, men dette gjelder ikke alle.

Konklusjon uten premisser

Det mest bemerkelsesverdige ved rapporten er dens diskusjonskapittel, som bare opptar én av i alt 37 sider. Diskusjonen sier intet om seleksjonsskjevhet, surrogatendepunkter, manglende kontrollgruppe eller at det muligens ble påvist flest CIN2+-tilfeller med mRNA-test. Når rapporten er bygd på antakelsene om at CIN2+ er et godt surrogatendepunkt og at gruppene er sammenliknbare, skulle man tro at mRNA-testen ble anbefalt. Tvert imot trekker man reservasjonsløst konklusjonen at det «anbefales at en bruker tester med bedre egenskaper enn» mRNA-testen.

Hva dreier konflikten seg om?

Konflikten rundt valg av HPV-test har vært både langvarig og betent. Den har sågar fått sindige gynekologer til å kreve at Helse-tilsynet gransker de sykehus som har brukt mRNA-testen (28). Forklaringen på det sterke engasjement kan, slik jeg ser det, neppe ligge i bekymring for kvinners helse. De som er opptatt av å forebygge flere krefttilfeller, har mye mer å hente hos de kvinner som ikke møter til screening, ikke blir etterkontrollert som anbefalt eller blir etterkontrollert uten HPV-test. Dersom man er opptatt av testegenskaper, burde man vurdere alle involverte tester, ikke bare HPV-testene. Hva vet vi om kvaliteten på celleprøver, histologiske undersøkelser,

kolposkopi og biopsitaking i Norge? Jeg kan ikke forstå denne saken uten å ta skjulte motiver i betraktning.

At jeg personlig har engasjert meg, er fordi jeg har sett meg lei på alle pseudoargumentene og mangelen på tallfesting av konsekvensene av forvaltningens vedtak. Jeg savner utredninger basert på etterprøvbare premisser, der fordeler så vel som ulemper ved screening løftes frem.

Konklusjon

Jeg vet ikke om DNA-tester er bedre enn RNA-tester i sekundærscreening for livmorkreft, og det er ingen stor sak i en større sammenheng. Som nevnt er andre forhold enn valg av test viktigere for det totale antall tilfeller av livmorkreft i Norge. Man ville blant annet tro at det er langt viktigere at nesten halvparten av kvinner med celleforandringer blir etterkontrollert uten HPV-test (4) enn hvilken test som brukes. Saken er likevel av betydning.

For det første reiser den spørsmål om kompetansen på screening og epidemiologi hos dem som skal forvalte screeningprogrammer, herunder screening for brystkreft og tarmkreft. Kan samfunnet stole på at forvaltningen er i stand til å overskue konsekvensene av screening når den tilsynelatende ikke har interesse av eller kompetanse til å tallfeste dem?

For det annet reiser den spørsmål om hvordan målgruppene skal informeres når den inviteres til screening. Hvorfor informerer ikke Kreftregisteret om antall falskt positive og falskt negative prøver eller om bivirkninger etter konisering? Er ikke slik informasjon like selvfølgelig som informasjon om legemiddelbivirkninger?

For det tredje tyder denne saken på at alle de falskt positive testresultatene ikke tillegges noen vekt. «Shared decision making» er blitt tema i mange land. «No decision about me without me,» sier den britiske statsminister (29). Vet virkelig legene hva som er best for kvinnene?

For det fjerde er det vanskelig å forstå at forvaltningen har brukt så mye tid, energi og prestisje på å få fjernet mRNA-testen dersom hensynet til kvinners helse var motivet. Med et slikt motiv finnes det bedre saker å gå løs på.

Jeg er positivt innstilt til screening, også for livmorkreft, som er kostnadseffektivt så langt vi kjenner til (30). Screening er imidlertid ønskelig bare så lenge fordelene for den enkelte kvinne er større enn ulem-

pene. Denne vurderingen varierer fra kvinne til kvinne, og ingen gynekologer eller helseøkonomer kan fortelle kvinnene hva som er best for dem. Det er på tide at helseministeren tar grep. Hun har mange muligheter: Flytting av det forvaltningsmessige ansvar for screening, oppnevning av nye rådgivningsgrupper, oppretting av uavhengige forskningsprogrammer, offentlig utredning – for å nevne noen. Hansken ligger hos helseministeren.

Ivar Sønbo Kristiansen (f. 1947)

er spesialist i samfunnsmedisin og professor ved Avdeling for helseledelse og helseøkonomi, Universitetet i Oslo. Han driver forskning innen helseøkonomi og risikokommunikasjon. Forfatter har fylt ut ICMJE-skjemaet og oppgir følgende interessekonflikter: Han har mottatt honorar fra mRNA-testprodusenten NORCHIP for foredrag om epidemiologi og testegenskaper i 2008 og 2009. Powerpoint-presentasjonene kan fås ved henvendelse til forfatteren. Han er medlem av gruppe II, som utreder bruk av HPV-testing i primærscreening for livmorkreft.

Litteratur

- Kreftregisteret. Masseundersøkelsen mot livmorkreft. www.kreftregisteret.no/no/Forebyggende/Masseundersokelsen-mot-livmorkreft/ (15.5.2012).
- Nilsen L. Vil fjerne refusjonstaksten for HPV-test. Dagens Medisin 2010; nr. 11: 4.
- Rindahl Endresen MJ, Andreassen T. Revidering og presisering av cervixscreeningprogrammet – Høring (brev av 11.2. 2011 fra Helsedirektoratet). Oslo: Helsedirektoratet, 2011.
- Haldorsen T, Skare GB, Bjørge T. Sekundærscreening med HPV-tester i Masseundersøkelsen mot livmorkreft. Oslo: Kreftregisteret, 2011.
- Rindahl Endresen MJ, Andreassen T. Vurdering av HPV-tester med hensyn på kravene i Masseundersøkelsen mot livmorkreft (brev av 20.12. 2011 Helsedirektoratet). Oslo: Helsedirektoratet, 2011.
- Skriftlig spørsmål fra Per Arne Olsen (FrP) til helse- og omsorgsministeren. www.stortinget.no/no/Saker-og-publikasjoner/Sporsmal/Skriftlig-sporsmal-og-svar/Skriftlig-sporsmal/?qid=52351 (15.5.2012).
- Sargent A, Bailey A, Almonte M et al. Prevalence of type-specific HPV infection by age and grade of cervical cytology: data from the ARTISTIC trial. Br J Cancer 2008; 98: 1704–9.
- Cancer in Norway 2009. Oslo: Kreftregisteret, 2011.
- Kristiansen IS. Hva er evidensbaserte råd? Tidsskr Nor Legeforen 2012; 132: 638–9.
- Johansen BK, Skare GB, Haldorsen T et al. Årsrapport 2008. Masseundersøkelser mot livmorkreft. Oslo: Kreftregisteret, 2009.
- Korfage IJ, van Ballegooijen M, Huvneers H et al. Anxiety and borderline PAP smear results. Eur J Cancer 2010; 46: 134–41.
- Bruinsma FJ, Quinn MA. The risk of preterm birth

following treatment for precancerous changes in the cervix: a systematic review and meta-analysis. BJOG 2011; 118: 1031–41.

- Albrechtsen S, Rasmussen S, Thoresen S et al. Pregnancy outcome in women before and after cervical conisation: population based cohort study. BMJ 2008; 337: a1343.
- Kreftregisteret. Celleprøven. <http://kreftregisteret.no/no/Forebyggende/Masseundersokelsen-mot-livmorkreft/Celleproven/> (18.5.2012).
- Burger EA, Kornør H, Klemp M et al. HPV mRNA tests for the detection of cervical intraepithelial neoplasia: a systematic review. Gynecol Oncol 2011; 120: 430–8.
- zur Hausen H. Papillomavirus and cancer: from basic studies to clinical application. Nat Rev Cancer 2002; 2: 342–50.
- ASCUS-LSIL Triage Study (ALTS) Group. Results of a randomized trial on the management of cytology interpretations of atypical squamous cells of undetermined significance. Am J Obstet Gynecol 2003; 188: 1383–92.
- Castle PE. The evolving definition of carcinogenic human papillomavirus. Infect Agent Cancer 2009; 4: 7.
- Meijer CJ, Berkhof J, Castle PE et al. Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women 30 years and older. Int J Cancer 2009; 124: 516–20.
- Kreftregisteret. Regnestykke på sensitivitet. www.kreftregisteret.no/no/Generelt/Nyheter/Regnestykke-pa-sensitivitet/ (3.4.2012).
- Tinacci G, Biggeri A, Pellegrini A et al. The use of digital images to evaluate the interobserver agreement on cervical smear readings in Italian cervical cancer screening. Cytopathology 2011; 22: 75–81.
- Elsheikh TM, Kirkpatrick JL, Fischer D et al. Does the time of day or weekday affect screening accuracy? A pilot correlation study with cytotechnologist workload and abnormal rate detection using the ThinPrep Imaging System. Cancer Cytopathol 2010; 118: 41–6.
- Stoler MH, Schiffman M. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. JAMA 2001; 285: 1500–5.
- Massad LS, Jeronimo J, Katki HA et al. The accuracy of colposcopic grading for detection of high-grade cervical intraepithelial neoplasia. J Low Genit Tract Dis 2009; 13: 137–44.
- Zuchna C, Hager M, Tringler B et al. Diagnostic accuracy of guided cervical biopsies: a prospective multicenter study comparing the histopathology of simultaneous biopsy and cone specimen. Am J Obstet Gynecol 2010; 203: 321. e1–6.
- Ostör AG. Natural history of cervical intraepithelial neoplasia: a critical review. Int J Gynecol Pathol 1993; 12: 186–92.
- McCredie MR, Sharples KJ, Paul C et al. Natural history of cervical neoplasia and risk of invasive cancer in women with cervical intraepithelial neoplasia 3: a retrospective cohort study. Lancet Oncol 2008; 9: 425–34.
- Bakke KA. Kreftalarm: Gynekologer krever granskning. Dagens Medisin 2012; nr. 4.
- Equity and excellence: liberating the NHS. London: Department of Health, 2010.
- Burger EA, Ortendahl JD, Sy S et al. Cost-effectiveness of cervical cancer screening with primary human papillomavirus testing in Norway. Br J Cancer 2012; 106: 1571–8.

Mottatt 1.5. 2012, første revisjon innsendt 15.5. 2012, godkjent 16.5. 2012. Medisinsk redaktør Siri Lunde.