

Tidlig diagnostikk av lungefibrose, emfysem og lungekreft blant nåværende og pensjonerte ansatte i norsk silisiumkarbidindustri, ved hjelp av MD-CT (multidetektor computertomografi).

Merete Drevvatne Bugge, Trond Mogens Aaløkken, Helge Kjuus

Nr. 23 (2008) Årgang 9 Serie: STAMI-rapport

ISSN nr. 1502-0932



Forord

Prosjektet som er omhandlet i denne rapporten ble initiert av styringsgruppen for prosjektet ”Eksposering og lungesykdommer i silisiumkarbidindustrien”, og ble gjennomført i tidsrommet fra høsten 2005 til våren 2007.

Vi vil gjerne takke alle som har bidratt til gjennomføring av prosjektet. Røntgensenteret Kristiansand og Capiro Røntgen i Trondheim har utført de fleste CT/røntgenundersøkelsene og en stor ekstrainsats ble utført av personalet for å få gjennomført undersøkelsene. Mange arbeidstimer er også lagt inn i prosjektet av bedriftshelsetjenestene ved silisiumkarbidverkene, representert ved Irene Espenes, Trond Klemetsen og Åse Sjøgård Sæterbakk. Dr. med. Kurt Hering ved Knappschaftskrankenhaus i Dortmund og Prof. Dr. John E. Parker ved West Virginia University Hospital har bidratt med tyding av henholdsvis CT- og røntgenbilder. Takk også til styringsgruppen, ved Vemund Digernes, Jon Efskind, Bjørn Erikson, Lars Petter Maltby og Bjørn Nordli.

Prosjektet har fått økonomisk støtte fra NHOs arbeidsmiljøfond, LO-forbundet Industri Energi og NHO-bransjeforeningen Norsk Industri.

Oslo, desember 2008
Merete Drevvatne Bugge
Trond Mogens Aaløkken
Helge Kjuus

Sammendrag

Bakgrunn

Computertomografi av lungene (CT) har de senere år fått en økende anvendelse i diagnostikk av lungesykdommer. Metoden er vist å være mer sensitiv og reproducerbar enn vanlig rtg.bilde av lungene, både for påvisning av støvlungesykdom, emfysem og lungekreft. Det er derfor reist spørsmål om CT-undersøkelse burde tilbys som screeningmetode til arbeidstakere i yrker med risiko for støvlungesykdom og lungekreft. Ansatte i norsk silisiumkarbidindustri er vist å ha økt risiko for lungekreft og andre lungesykdommer. Formålet med dette prosjektet har vært å sammenlikne multidetektor computertomografi (MD-CT) undersøkelse av lungene med andre metoder for vurdering av lungesykdom blant ansatte i SiC-industrien. Vi har videre ønsket å skaffe oss erfaring med gjennomførbarheten av MD-CT som screeningmetode ved helseundersøkelse av støvutsatte yrkesgrupper.

Materiale og metode

Blant 413 inviterte nåværende og tidligere ansatte i norsk silisiumkarbidindustri ble 275 menn undersøkt med multidetektor computertomografi (MD-CT) og røntgenundersøkelse av lungene (responsrate: 67 %). I tillegg ble det utført lungefunksjonsundersøkelse (spirometri) på 173 av disse. 2 CT-spesialister har gjennomgått MD-CT-undersøkelsene, og 2 ILO-sertifiserte B-readere har gjennomgått røntgenbildene.

Resultater

MD-CT undersøkelsen ga svært mange positive funn. Kun 6 % av filmene ble av begge tydere klassifisert som "negativ film", dvs. helt uten funn. Tilsvarende andel "negativ film" basert på røntgenbildene var 49 %. Pneumokoniose ble påvist av begge tydere i 23 av CT-bildene og bare ett av rtg. thorax-bildene. I fem av tilfellene anga begge tydere funn med mistanke om lungekreft. Samsvaret mellom CT-tyderne, mellom CT- og røntgenundersøkelsene og mellom CT- og spirometriundersøkelsene var relativt lavt, spesielt med hensyn til klassifisering av positive funn (tegn på sykdom). Etter primær CT vurdering av en av tyderne ble i alt 80 personer (30 %) fulgt opp med tilleggundersøkelser. Vi er ikke kjent med resultatet fra oppfølgingsundersøkelsene.

Konklusjon

CT-undersøkelsen er en svært sensitiv metode, med et høyt antall positive funn. En stor andel av de aktuelle funn som krevde oppfølging var trolig uten klinisk relevans, og kan ha medført en belastning for den enkelte. Videre var sikkerheten i diagnosen ved CT-undersøkelse alene relativt lav, i og med at samsvaret mellom CT-tyderne var lav. Dersom et screeningprosjekt med CT-undersøkelse planlegges, må det i tillegg til selve undersøkelsen legges inn planer for og kostnader ved den direkte oppfølgningen av positive funn. Inntil resultater fra store randomiserte studier foreligger, kan vi foreløpig ikke på faglig grunnlag tilrå lungekreftscreening ved bruk av CT i yrkeseksponerte høyrisikogrupper. CT anbefales imidlertid ved individuell oppfølging av usikre røntgenfunn og ved lungesympomer.

Innledning / bakgrunn

Helseovervåkning av støvutsatte arbeidstakere

Ifølge norsk lov (asbestforskriften¹ / kjemikalieforskriften²) skal arbeidstakere som utsettes for kjemisk påvirkning få tilbud om helseundersøkelse jevnlig. For arbeidstakere som i sitt arbeid er utsatt for mye støv har en slik helseundersøkelse vanligvis blant annet medført røntgenundersøkelse av lungene med noen års mellomrom.

Ansatte i norsk silisiumkarbidindustri har helt siden 1940-tallet, som et ledd i rutinemessige helsekontroller blitt undersøkt regelmessig med røntgen/skjerm bilde av lungene. Undersøkelsen har som regel funnet sted ca. hvert tredje år. Formålet med undersøkelsen har vært å oppdage tidlige tegn på pneumokoniose (støvlungesykdom), som pleuraplakk (fortykkelse av brysthinnen) og fortetninger i lungevevet (fibrose).

Kreft- og dødelighetsstudier av ansatte i norsk silisiumkarbidindustri, utført i regi av Kreftregisteret i samarbeid med de lokale bedriftshelsetjenester, har påvist en økt forekomst av lungesykdom sammenlignet med normalbefolkningen, både når det gjelder lungekreft, lungefibrose og kronisk obstruktive lungesykdommer (KOLS)³⁻⁵.

De senere år er det kommet undersøkelsesmetoder som har vist seg langt mer sensitive enn røntgen når det gjelder tidlig diagnostikk av lungefibrose. Computertomografi, CT, er sammenlignet med røntgenfotografering av lungene i flere studier, og har vist seg mer sensitiv og mer reproducerbar (høyere inter-reader og intra-reader samsvar) både når det gjelder påvisning av asbestose⁶, silikose⁷ og emfysem^{8,9}, samt ved tidlig diagnostikk av lungekreft¹⁰. KOLS, med unntak av emfysem, kan ikke oppdages ved hjelp av denne metoden.

De gode resultatene disse studiene har gitt, har gjort at det er reist spørsmål ved om CT-undersøkelse burde tilbys som screeningmetode, både til befolkningsgrupper med høy lungekreftrisiko, spesielt storryktere, og, i stedet for røntgen thorax, til arbeidstakere i yrker med risiko for pneumokoniose og lungekreft.

To studier (Bergin⁷ og Remy-Jardin¹¹) har sammenlignet CT og røntgen ved screening for pneumokoniose. Selv om begge fant at CT var en mer pålitelig undersøkelsesmetode, konkluderte man i begge studier med at kostnadene ved CT-undersøkelse var så betydelige sammenlignet med røntgen, og effektene så begrensede at CT-undersøkelse som fast screeningmetode ikke kunne anbefales. Imidlertid anbefalte man videre utredning med CT dersom det ved røntgenundersøkelse var funn som ga mistanke om pneumokoniose.

Lungekreftscreening

Lungekreft er en av de vanligste formene for kreft, og en sykdom som har svært lav overlevelse hvis den oppdages etter at den gir kliniske symptomer. Det å oppdage lungekreft i et tidlig stadium, og dermed kunne operere og fjerne kreftsvulsten før den har spredt seg til andre organer, har i mange år, og av mange miljøer, vært ansett som en mulighet til å redde svært mange liv.

Allerede på 1960- og -70-tallet ble det igangsatt screeningprosjekter med tanke på tidligdiagnostikk av lungekreft¹². Metodene som først og fremst ble benyttet var røntgen thorax og sputumundersøkelse (undersøkelse av celler i spyttprøve). Studiene ga primært lovende resultater, i og med at det ble oppdaget flere tilfeller av lungekreft i et tidlig stadium i screeninggruppene enn i kontrollgruppene. Imidlertid har det vist seg i oppfølgingsundersøkelser at denne økte diagnostikken av tidlig lungekreft ikke ga noen redusert forekomst av lungekreft i sent stadium i screeninggruppene, og medførte heller ikke noen redusert samlet dødelighet i forhold til i kontrollgruppene¹³. Det ble derfor, på bakgrunn av disse studiene, ikke funnet grunnlag for å anbefale lungekreftscreening ved hjelp av disse metodene.¹⁴

På den annen side hadde forsøk med en ny billedteknologi, såkalt spiral CT-teknologi gitt lovende resultater, og det ble derfor på slutten av 1990-tallet startet opp pilotforsøk i USA, med bruk av lav-dose CT som screeningmetode i befolkningsundersøkelser blant storryktere^{10;15}.

I Norge ble det i 2002 tatt initiativ til planlegging av en større nasjonal befolkningsundersøkelse med CT- og sputumundersøkelse, hvor man i tillegg til storryktere ønsket å inkludere personer med yrkeseksponering fra industri med forhøyet lungekreftisiko, som asbesteksponerte fra Norcem på Slemmestad, nikkeleksponerte fra Falconbridge og personer med eksponering fra silisiumkarbidindustrien. Studien skulle gå over flere år med årlige CT- og sputumundersøkelser, og formålet med studien skulle være å utrede hvorvidt en slik screening ville påvirke lungekreftdødelighet i populasjonen. Videre ønsket man å vurdere om det at man kunne oppdage og behandle andre lunge- og hjerte-karsykdommer på et tidligere tidspunkt kunne gi en reduksjon i dødelighet av andre årsaker.

Det lyktes ikke å skaffe tilstrekkelig finansiering til denne planlagte nasjonale lungekreftstudien, og den måtte derfor skrinlegges. Imidlertid var det, både fra prosjektledelsen og fra styringsgruppen i det allerede pågående silisiumkarbidprosjektet, ønske om å gjøre en studie i mindre skala i silisiumkarbidindustrien, hvor bruk av CT som screeningverktøy kunne sammenlignes med røntgen thorax og spirometri. Man ønsket derfor å vurdere gjennomførbarheten av CT-screening under de eksisterende forhold, samt vurdere kost-effektivitetsforhold knyttet til et slikt screeningtiltak.

Målsetning med prosjektet

1. Sammenlikne multidetektor computertomografi (MD-CT) undersøkelse av lungene med andre metoder for vurdering av lungesykdom blant ansatte i SiC-industrien:
 - a) Sammenligne forekomst av pneumokoniose (fibrose / pleuraplakk) og lungekreft ved MD-CT-undersøkelse med tilsvarende forekomst ved røntgenundersøkelse. Sammenligne forekomst av emfysem ved MD-CT-undersøkelse med emfyseparametre fra spirometriundersøkelse.
 - b) Sammenlikne resultatet av tydingen av hhv. CT- og rtg.bilder foretatt av to uavhengige ekspert-tydere (inter-reader samsvar)
 - c) Ut fra det ovenstående vurdere om MD-CT bidrar til en vesentlig forbedring av tidligdiagnostikk i forhold til de andre undersøkelsene.
 - d) Vurdere om resultater fra MD-CT-undersøkelse i større grad enn andre metoder gir informasjon som belyser eksponerings-responsammenhenger i SiC-industrien (sammenholde resultater med eksponeringsopplysninger fra jobbeksponeeringsmatrise).
2. Vi har videre ønsket å skaffe oss erfaring med gjennomførbarheten av MD-CT som screeningmetode ved helseundersøkelse av støvutsatte yrkesgrupper.
3. Vi har i tillegg ønsket å foreta en kost-effektivitetsvurdering av MD-CT som screeningmetode sammenlignet med tradisjonelle metoder brukt i forbindelse med helseundersøkelse av støvutsatte ansatte.

Materiale og metode

Multidetektor computertomografi (MD-CT) tilhører den siste generasjonen CT-utstyr. På veldig kort tid (5-10 sekunder) kan denne maskinen scanne lungene i detalj, og filmen som tas opp gir mulighet for både detaljert screening av pneumokoniose, og oversiktsscreening med tanke på kreft. Tidligere krevet disse to formålene to forskjellige undersøkelser (HR-CT for pneumokoniose, spiral CT for kreft).

Stråledosen ved MD-CT er også lavere enn ved de tidligere CT-teknikkene. Dosenivået ved en MD-CT-undersøkelse er ca. 1 mSv. Dette er ca. 10 ganger høyere enn ved en vanlig røntgen thorax, og omtrent på samme nivå som ved en mammografiundersøkelse.

Inklusjonskriterier for studien var alder 50-75 år og minst 5 års ansettelse ved et av de tre norske silisiumkarbidverkene. Det er for få kvinner ansatt ved verkene til at man kunne inkludere kvinner i den vitenskapelige delen av denne studien, men etter ønske fra bedriftene fikk også kvinner som fylte inklusjonskriteriene tilbud om samme undersøkelse som mennene. I våre analyser er kun menn inkludert.

Ved gjennomgang av ansettelseslister ble det funnet at 413 menn fylte inklusjonskriteriene. Av disse var 157 fortsatt i arbeid i industrien, mens de resterende 256 hadde sluttet i arbeidet i SiC-industrien, enten fordi de hadde gått over i annet arbeid, eller fordi de var pensjonister. Selve undersøkelsen ble organisert noe forskjellig avhengig av verk og ansettelsesforhold. De 130 som fortsatt arbeidet ved et av SiC-verkene på Sørlandet, ble kontaktet av bedriftshelsetjenesten ved verket, og fikk informasjon om prosjektet i forbindelse med en konsultasjon. De fikk utdelt et informasjonsbrev, og ble, dersom de ønsket å delta i studien, bedt om å fylle ut et samtykkeskjema og et spørreskjema som ble levert inn med det samme. I spørreskjemaet ble deltagerne bedt om å redegjøre for støveksponering utenom silisiumkarbidindustrien, kjent hjerte- og/eller lungesykdom, og røykehistorie. Deltagerne ble også undersøkt med spirometri (lungesfunksjonsundersøkelse). Bedriftshelsetjenesten hadde direkte kontakt med Røntgensenteret i Kristiansand, og gjorde avtale om undersøkelsestidspunkt, og organiserte også felles transport til Røntgensenteret for deltagerne. De øvrige, 27 personer ansatt ved smelteverket i Trøndelag, samt de 256 som hadde sluttet i SiC-industrien, fikk tilsendt informasjonsbrev om studien, samtykkeskjema og spørreskjema til sin privatadresse. De som returnerte utfylt samtykkeskjema og spørreskjema ble henvist til CT- og røntgenundersøkelse av prosjektlegen.

Til sammen 301 personer undertegnet samtykke til studien, og leverte inn spørreskjema, 3 avslo deltagelse, og 107 personer svarte ikke på henvendelsen, etter en purring. Av de 301 som samtykket til deltagelse var 121 i arbeid ved et av verkene på Sørlandet, og ble via bedriftshelsetjenesten henvist til MD-CT-undersøkelse og røntgen thorax ved Røntgensenteret i Kristiansand. I alt 25 personer ansatt ved SiC-verket i Trøndelag, samt 155 personer som hadde sluttet i arbeidet i silisiumkarbidindustrien, ble henvist til MD-CT-undersøkelse og røntgen thorax av prosjektlegen. I tillegg ble deltagerne oppfordret til å kontakte

bedriftshelsetjenesten ved SiC-verket for å utføre en lungefunksjonsundersøkelse. De aller fleste ble henvist til MD-CT/røntgen enten ved Røntgensenteret i Kristiansand eller Capio Røntgen i Trondheim, ettersom de fleste fortsatt var bosatt i nærheten av SiC-verkene i Arendal, Lillesand eller Orkanger. Noen personer hadde imidlertid flyttet til andre deler av landet, så det ble i tillegg utført undersøkelser ved Rikshospitalet, Helse Stavanger HF, Nordlandssykehuset HF og Sykehuset Østfold Fredrikstad.

Av 301 personer henvist til MD-CT/røntgenundersøkelse, møtte 275 opp til undersøkelsen. Det ble i tillegg utført spirometrisk undersøkelse på 173 av disse. Alle undersøkelser var gratis for deltagerne, og reiseutgifter ble refundert.

Undersøkelsene ble gjennomført i perioden 14. desember 2005 til 22. mai 2006. Røntgenavdelingene ble bedt om å melde fra til prosjektlegen dersom det ved gjennomgang av CT-filmen ble gjort funn som burde følges opp umiddelbart. Forøvrig ble det ved røntgenavdelingene ikke gjort en ordinær beskrivelse av CT-filmene eller røntgenbildene. Det ble meldt fra om slike funn hos 6 personer, og MD-CT-filmene for disse ble umiddelbart oversendt til tydning hos radiolog (TMA). To personer ble etter tydningen anbefalt henvist til utredning ved lungeavdeling, og dette ble raskt fulgt opp.

Det ble laget digitale kopier av alle MD-CT-filmene og røntgenbildene. MD-CT-filmene ble oversendt til to radiologer for tydning (TMA og KGH), mens røntgenbildene ble ILO-klassifisert av to ILO-sertifiserte B-readere (JEP og MDB). Ved klassifikasjon av røntgenbildene ble det benyttet et standardisert klassifikasjonsskjema for registrering av funn forenlige med pneumokoniose. Dette skjemaet har også rubrikker for avkrysning ved andre funn, som emfysem, noder, mistanke om cancer, etc. Det foreligger ikke et tilsvarende standardisert skjema for registrering av CT-funn, og vi valgte derfor å ta utgangspunkt i et forslag til klassifiseringsskjema for registrering av pneumokoniose ved CT, som er utarbeidet av en internasjonal gruppe.¹⁶ Dette skjemaet har også rubrikker for registrering av andre funn, i likhet med røntgenskjemaet. I tillegg ønsket vi en nærmere beskrivelse av eventuelle CT-påviste noder, så vi utarbeidet et tilleggsskjema i samarbeid med radiologene, hvor størrelse, lokalisasjon og nærmere beskrivelse av eventuelle noder kunne registreres, i tillegg til en anbefalt oppfølging av disse funnene.

Etter at klassifisering av CT-filmer og røntgenbilder er utført av to CT-eksperter og to ILO-sertifiserte B-readere, ble registreringer av alle skjemaer lagt inn i regnearkprogrammet Excel, og videre analyser er utført i statistikkprogrammet STATA.

For samsvarsundersøkelsen mellom to tydere er det beregnet *Observert samsvar* (også kalt simple agreement / raw agreement), som er andelen av klassifiseringer som tyderne er enige om. *Positivt spesifikt samsvar* og *negativt spesifikt samsvar* er andelen klassifiseringer i de gitte kategorier (positiv / negativ) som tyderne er enige om. *Kappa* er et uttrykk for forholdet mellom det observerte samsvar og det samsvar man ville forvente på grunnlag av tilfeldighet. Videre utdyping av disse begrepene er omtalt av Hripcsak et al.¹⁷

Resultater

Videre oppfølging av funn

Etter at den ene radiologen hadde fullført tydingen av CT-filmene, i oktober 2006, ble anbefalt oppfølging av funn utført. Til sammen ble 80 av 275 personer (29%) på grunnlag av denne første tydingen funnet å ha behov for en oppfølging av funn. Funnene var imidlertid av svært variert alvorlighetsgrad, og det ble satt opp en prioriteringsrekkefølge, både med tanke på prosjektlegens henvisninger, og med tanke på prioritering ved den lungeavdelingen som fikk tilsendt de fleste henvisningene. Denne avdelingen fikk i alt 34 henvisninger, fordelt på to grupper (1. og 2. prioritet). De personene som hadde funn som ble vurdert til å være av lavest alvorlighetsgrad, ble kun anbefalt å kontakte sin fastlege for videre vurdering. Tabell 1 gir en oversikt over de funn som ble gjort, og som ble vurdert til å ha behov for en eller annen form for oppfølging.

Tabell 1. Oversikt over type funn og oppfølging blant 80 personer^a med CT-forandringer

Funn	Gradering	Antall funn	Oppfølging
Pneumokoniose	>8	12	Henvisning lungeavdeling
Emfysem	>8	14	Henvisning lungeavdeling
Nodul	>4 mm / multiple	24	Henvist lungeavdeling / Kontroll CT
	<4 mm, røker	18	Anbefalt kontroll-CT i regi av fastlegen
	<4 mm, ikke-røker	9	Ingen videre oppfølging*
Aortaaneurysme / ektasi		6	Henvist karkirurgisk avdeling for oppfølging
Ekstensiv coronarkalk		17	Orientering, anbefalt kontakt med fastlegen for vurdering av oppfølging
Annet (lymfeknuter, mistanke om infeksjon, RA)		8	Henvisning lungeavdeling

^aEn person i denne oversikten kan ha flere forskjellige funn

*Gradering av og valg av oppfølging ved noder er gjort i samsvar med anbefalinger fra *the Fleischner Society*¹⁸

Antallet funn som måtte følges opp, viste seg å være svært høyt, og selv om avtale var gjort i forkant med den sykehusavdelingen som måtte antas å få den største delen av henvisningene, ble det vanskelig for denne avdelingen å følge opp alle henvisningene i løpet av den tiden de hadde stipulert. Spesielt viste det seg at det her oppsto vanskeligheter med få utført kontroll-CT ved egen røntgenavdeling på grunn av lange ventelister, og bl.a. dette forhold gjorde at oppfølgingsundersøkelsene strakte seg over lang tid, med de siste undersøkelsene gjennomført i løpet av sommeren 2007.

Samtykkeerklæringen som ble innhentet ved starten av prosjektet dekker analyse av samsvar mellom resultater fra de forskjellige undersøkelsene (CT, rtg og spirometri), samt en vurdering av funn i relasjon til eksponering. Vi innhentet imidlertid ikke samtykke til å gjøre vitenskapelige analyser av den videre oppfølgingen av funnene, og av de endelige resultatene av undersøkelsene. Vi har derfor ikke grunnlag for her å gjøre en vurdering av relevansen av de funnene som ble gjort.

Sammenlikning av røntgen- og CT-funn.

Tabell 2 oppsummerer de viktigste funn hos alle 4 lesere. Leser A og B er CT-tyderne, mens leser C og D er de ILO-sertifiserte B-readerne av røntgenbildene. Leser A, C og D har klassifisert alle 275 filmene, mens det fra leser B kun er mottatt klassifisering av 270 av filmene. Antall funn er langt høyere ved CT-undersøkelsen enn ved røntgen, noe som fremkommer tydelig i kategorien ”negativ film”. CT-tyderne klassifiserer filmen som helt negativ (ingen funn) i henholdsvis 40 (15 %) og 33 (12%) av tilfellene, mens de tilsvarende tall for røntgenbildene er 179 (65%) og 164 (60%). Dette gjenspeiler seg i alle kategorier av funn, samtidig som det for enkelte typer funn er en viss variasjon også mellom leserne. Størst forskjell mellom CT-tyderne er det i kategoriene ”store pneumokoniotiske fortetninger” (hhv. 34 og 8 positive funn) og ”noduler” (hhv. 59 og 92 positive funn). Man kan tenke seg at noduler og store pneumokoniotiske fortetninger kan ha visse likhetstrekk, og at leserne kan ha en tendens til å klassifisere disse i hver sin kategori. Vi har derfor valgt å legge inn en ekstra kategori, hvor disse to er slått sammen, og forskjellen mellom leserne blir da mindre (hhv. 81 og 92 positive funn). Ingen av røntgentyderne har klassifisert noe funn som *store pneumokoniotiske fortetninger*. Mistanke om lungekreft ble av de to CT-lesere angitt i hhv. 8 og 12 tilfelle (totalt 15 tilfelle), og i 3 tilfelle av en av røntgenleserne.

Tabell 2: Funn ved henholdsvis CT- og røntgenundersøkelse, alle 4 lesere

	Leser A (CT)	Leser B (CT)	Leser C (røntgen)	Leser D (røntgen)
Antall filmer lest	275	270	275	275
Negativ film	40	33	179	164
Pneumokoniose	51	61	5	2
Store pneumokon. fortetninger*	34	8	0	0
Pleuraplakk	73	76	11	18
Pleurakalk	22	27	6	5
Emfysem	62	59	2	15
Noduler*	59	92	3	8
Nodul / store fortetninger*	81	92	3	8
Mistanke om kreft	8	12	0	3

* Ettersom leserne til en viss grad har krysset av for både ”store pneumokoniotiske fortetninger” og ”noduler” på samme skjema, er summen av disse to høyere enn den sammenslåtte verdien.

Tabell 3: Funn og samsvar mellom lesere. CT

	N	A1/B1	A1/B0	A0/B1	A0/B0	Observed agreement	Pos. specif. agreem.	Neg. specif. agreem.	Kappa
Negativ film	270	17	23	16	214	0,86	0,47	0,92	0,38
Pneumokoniose	270	23	27	38	182	0,76	0,41	0,85	0,26
Runde fortetninger	270	1	17	0	252	0,94	0,11	0,97	0,10
Irregulære fortetn.	270	18	17	42	193	0,78	0,38	0,87	0,26
Store fortetninger	270	5	29	3	233	0,88	0,24	0,94	0,20
Pleuraplakk	270	52	17	24	177	0,85	0,72	0,90	0,61
Pleurakalk	270	17	4	10	239	0,95	0,71	0,97	0,68
Emfysem	270	44	18	15	193	0,88	0,73	0,92	0,65
Tuberkulose	270	0	0	14	256	0,95	0,00	0,97	0,00
Nodul	270	41	17	51	161	0,75	0,55	0,83	0,38
Mistanke om kreft	270	5	3	7	255	0,96	0,50	0,98	0,48
Behov for oppfølging	270	16	12	41	201	0,80	0,38	0,88	0,28
Store fortetninger / noduler	270	49	31	43	147	0,73	0,57	0,80	0,37

Tabell 3 viser registrerte funn og samsvar mellom de to leserne av CT-filmene. Av 270 filmer som ble lest av begge lesere, ble 40 klassifisert som uten positive funn (negativ film) av leser A, og 33 ble klassifisert som negativ av leser B. Det er imidlertid bare 17 disse filmene som begge lesere er enige om er negative, hvilket innebærer at hos 253 av 270 deltagere (94 %) i studien ble det registrert funn av en eller begge lesere. Leser A har registrert funn forenlig med pneumokoniose i 50 filmer, og leser B har registrert pneumokoniotiske forandringer hos 61 deltagere. I 23 filmer har begge lesere registrert pneumokoniose. Dette gir et observert samsvar på 0,76, et positivt spesifikt samsvar på 0,41, et negativt spesifikt samsvar på 0,85, og kappa på 0,26. Samsvar mellom leserne varierer avhengig av type funn. Det er et relativt høyt samsvar ("Substantial agreement" – se Landis og Koch¹⁹) for registrering av pleuraplakk ($\kappa=0,61$), pleurakalk ($\kappa=0,68$) og emfysem ($\kappa=0,65$), mens når det gjelder registrering av noduler, og vurdering av om filmen er negativ, er samsvaret lavere (Kappa = 0,38 – "Fair agreement"). Funn forenlig med pneumokoniose har et enda lavere samsvar. For tuberkulose er kappa på 0,00, dette har sammenheng med at det er bare en av leserne som har registrert funn forenlig med tuberkulose.

Jevnt over ligger *negativt spesifikt samsvar* på 0,8-0,9, altså at leserne er godt enige om hvilke filmer som ikke har det aktuelle funnet. Det er det *positive spesifikke samsvar* som gir lave resultater.

Om samsvar og kappverdier ligger relativt lavt ved CT-undersøkelsen, er samsvaret ikke uventet enda dårligere ved røntgenklassifiseringen.

Tabell 4: Funn og samsvar mellom lesere. Røntgen

	N	C1/D1	C1/D0	C0/D1	C0/D0	Observed agreement	Pos. specif. agreem.	Neg. specif. agreem.	Kappa
Negativ film	270	132	47	32	64	0,71	0,77	0,62	0,39
Pneumokoniose	270	1	4	1	269	0,98	0,29	0,99	0,28
Pleuraplakk	270	7	4	11	253	0,95	0,48	0,97	0,46
Pleurakalk	270	3	3	2	267	0,98	0,55	0,99	0,54
Emfysem	270	1	1	14	259	0,95	0,12	0,97	0,11
Tuberkulose	270	0	0	2	273	0,99	0,00	1,00	0,00
Nodul	270	0	3	8	264	0,96	0,00	0,98	-0,02
Mistanke om lungekreft	270	0	0	3	272	0,99	0,00	0,99	0,00

Kategoriene plakk og pleurakalk gir høyest samsvar, med kappa på hhv. 0,46 og 0,54, mens andre kategorier gir et lavere samsvar. I alt 132 røntgenbilder (49 %) er klassifisert som helt negative av begge B-readerne.

Tabell 5: Funn og samsvar mellom metoder. CT og røntgen

Rtg thorax - CT. 2x2 lesere. Samsvar der begge lesere er enige om positive funn

	N	R1/C1	R1/C0	R0/C1	R0/C0	Observed agreement	Pos. specif. agreem.	Neg. specif. agreem.	Kappa
Negativ film	270	9	121	8	132	0,52	0,12	0,67	0,01
Pneumokoniose	270	1	0	22	247	0,92	0,08	0,96	0,08
Pleuraplakk	270	6	0	46	218	0,83	0,21	0,90	0,17
Pleurakalk	270	2	0	15	253	0,94	0,21	0,97	0,20
Emfysem	270	1	0	43	226	0,84	0,04	0,91	0,04
Nodul	270	0	0	41	229	0,85	0,00	0,92	0,00
Mistanke om lungekreft	270	0	0	5	265	0,98	0,00	0,99	0,00

Tabell 5 viser samsvar mellom CT-undersøkelse og røntgenundersøkelse, når positivt CT-funn betyr at begge CT-tydere er enige om funnet, og positivt røntgenfunn betyr at begge røntgenlesere er enige. Det er ingen tilfeller av noduler eller cancermistanke hvor alle 4 lesere er enige om funnet. Alle 4 lesere har vært enige om registrering av ett tilfelle av pneumokoniose, 6 tilfeller av pleuraplakk, 2 av pleurakalk, og ett tilfelle av emfysem. Ved en sammenligning med tabell 4 finner vi at i nesten alle tilfeller av funn (med unntak av ett tilfelle av pleuraplakk og ett tilfelle av pleurakalk) hvor begge røntgenlesere har vært enige, er dette funnet også registrert av begge CT-tydere. Det er en hel rekke funn hvor begge CT-

leserne har vært enige, og hvor dette funnet ikke er registrert av begge røntgenlesere (Kolonne R0/C1). Dette gjelder blant annet 41 registreringer av noder og 5 registreringer av mistanke om cancer. Det er ingen tilfeller av det motsatte (R1/C0). Kappa er lav for samsvar mellom CT og røntgen når det gjelder alle typer funn.

Tabell 6: Funn og samsvar mellom metoder. CT og røntgen

Rtg thorax - CT. 2x2 lesere. Samsvar der minst en av leserne melder om positive funn

	N	R1/C1	R1/C0	R0/C1	R0/C0	Observed agreement	Pos. specif. agreem.	Neg. specif. agreem.	Kappa
Negativ film	270	52	155	4	59	0,41	0,40	0,43	0,10
Pneumokoniose	270	4	2	84	180	0,68	0,09	0,81	0,05
Pleuraplakk	270	16	5	77	172	0,70	0,28	0,81	0,18
Pleurakalk	270	6	1	25	238	0,90	0,32	0,95	0,29
Emfysem	270	10	6	67	187	0,73	0,22	0,84	0,13
Nodul	270	8	3	101	158	0,61	0,13	0,75	0,06
Mistanke om lungekreft	270	1	2	14	253	0,94	0,11	0,97	0,09

Tabell 6 viser samsvar mellom CT-funn og røntgenfunn, der positivt CT-funn betyr at minst én av CT-tyderne har registrert funnet, og tilsvarende for røntgen. I 8 tilfeller har minst én røntgenleser og minst én CT-tyder registrert nodul, og i ett tilfelle har minst én røntgenleser og én CT-tyder registrert mistanke om cancer. Det er litt høyere samsvar mellom CT og røntgen enn i tabell 5, men også her er samsvaret svært begrenset.

Mellom CT og spirometri er det gjort 4 samsvarsvurderinger, avhengig av hvor man definerer grensen for et positivt utfall (ved spirometri er "positivt utfall" KOLS, ved CT er "positivt utfall" emfysem – det kan diskuteres om disse er sammenlignbare størrelser, men det er det nærmeste vi kommer):

- Grensen for KOLS settes ved "mild KOLS" = $FVC/FEV1 < 70\%$, $FEV1 > 80\%$ av forventet verdi²⁰
- Grensen for KOLS settes ved "moderat KOLS" = $FVC/FEV1 < 70\%$, $FEV1$ er mellom 50 og 80 % av forventet
- Kravet for CT-definert emfysem er at begge leserne har notert emfysem
- Kravet for CT-definert emfysem er at én av leserne har notert emfysem

Tabell 7: Samsvar mellom KOLS-parametere ved spirometri og notert "emfysem" ved CT

	N	KOLS / Emfysem	KOLS / ikke emfysem	Ikke KOLS / Emfysem	Ikke KOLS / ikke emfysem	Observed agreement	Positive spec. agreement	Negative spec. agreement	Kappa
a / c	167	10	14	15	128	0,83	0,41	0,90	0,31
a / d	167	14	10	32	111	0,75	0,40	0,84	0,26
b / c	167	7	8	18	134	0,84	0,35	0,91	0,27
b / d	167	10	5	36	116	0,75	0,33	0,85	0,22

I tabell 7 presenteres samsvaret mellom resultatene fra spirometriundersøkelsene og hvorvidt det er klassifisert ”emfysem” ved CT-lesningen. CT gir noe flere ”funn” enn spirometri, om enn ikke så uttalt som ved sammenligningen med røntgen. Samsvaret mellom metodene ligger, uavhengig av valg av definisjon (a/b – c/d), på et lavt nivå, på grunn av et lavt positivt spesifikt samsvar. Valg av definisjon (a/b – c/d) har lite å si for resultatet av samsvarsanalysen. Kappa ligger på mellom 0,22 og 0,31, ifølge Landis og Koch¹⁹ ”Fair agreement”.

Diskusjon

Denne undersøkelsen viser at MD-CT er en svært sensitiv metode for påvisning av røntgenologiske lungeforandringer i en støvekspontert mannlig befolkning mellom 50 og 75 år. I alt 253 av 270 CT-bilder (94 %) ble klassifisert som ”positivt funn” av minst en av tyderne. Men selv på vanlig rgt. thorax ble 51 % av bildene beskrevet med ”positive funn”.

Dersom vi legger enighet mellom de to CT-lesere til grunn, ble det blant 270 nåværende og tidligere SiC-arbeidere påvist 23 tilfelle av pneumokoniose, 52 tilfelle av pleuraplakk, 44 tilfelle av emfysem og fem tilfelle med mistanke om lungekreft. Tilsvarende tall for samsvar blant rgt.leserne var ett tilfelle av pneumokoniose, syv tilfelle av pleuraplakk, ett tilfelle av emfysem og ingen tilfelle av mistenkt lungekreft. Dette funn illustrerer at sensitiviteten for positive funn ved MD-CD er dramatisk mye høyere enn for rgt.thorax. Computertomografi av lungene anses derfor som en nødvendig oppfølgingsundersøkelse ved uklare røntgenologiske lungefunn.

Det høye antallet positive funn påvist i dette prosjektet, innebærer imidlertid at en stor andel friske personer, med ”funn” uten klinisk relevans må følges opp med videre undersøkelser uten at det foreligger sykdom. Videre er det en relativt stor uenighet mellom erfarne tydere m.h.t. hva som skal klassifiseres som ”positive funn”. Disse forhold problematiserer bruken av MD-CT som screeningmetode. En screeningundersøkelse ved hjelp av MD-CT vil kreve store ressurser i forbindelse med oppfølging og videre utredning av den enkelte pasient. Kapasiteten innenfor det norske offentlige helsevesenet er pr. i dag for begrenset til at dette vil kunne ivaretas her. I tillegg til kostnader ved selve CT-undersøkelsen bør også kostnader ved anvendelse av spesialkompetanse til tyding av filmer og oppfølging av funn innberegnes.

Erfaringer med CT-prosjektet i silisiumkarbidindustrien

Den foreliggende undersøkelsen som nå er gjennomført i silisiumkarbidindustrien har gitt oss viktige erfaringer relatert til en eventuell gjennomføring av screening av lungene ved hjelp av computertomografi.

For det første: som forventet gir en CT-undersøkelse langt flere funn, og tydeligere funn, enn en røntgenundersøkelse. Dette er en stor styrke ved CT i forbindelse med utredning av pasienter, hvor symptomer, og uklare funn ved røntgen, kan få en sikrere avklaring ved hjelp av CT. I screeningsammenheng er det noe mer usikkert om CTs evne til å vise mange forandringer gir fordeler som oppveier ulempene. Den umiddelbare fordelene, og grunnen til at CT blir vurdert som screeningredskap, er at man har mulighet til å oppdage sykdommer tidlig og behandle dem tidlig. Samtidig er det en betydelig ulempe at man i tillegg oppdager svært mange forandringer i lungene som ikke har klinisk relevans, men som er umulig å skille fra alvorlig sykdom bare ved hjelp av den ene CT-undersøkelsen. En CT-screening vil derfor innebære at svært mange mennesker må utredes videre etter CT-undersøkelsen, uten at de feiler noe.

Dette fører oss videre til erfaring nr. to: Det norske offentlige helsevesenet av i dag er ikke dimensjonert for å ivareta oppfølgingen etter en slik screeningundersøkelse. Vi hadde i vårt prosjekt spesiell kontakt med en lungeavdeling på et sykehus, og samarbeidet der med en lungelege som var interessert i og ønsket å følge opp disse pasientene. Avdelingen fikk henvist 34 personer til oppfølging, og det tok over ½ år før den siste hadde vært inne til undersøkelse. Hovedårsaken til at det tok så lang tid, viste seg å være kapasitetsproblemer på røntgenavdelingen på sykehuset, hvor kontroll-CT skulle utføres. Ved et annet sykehus fikk vi beskjed om at funn oppdaget ved screening ikke kunne prioriteres ved lungeavdelingen, og noen av disse pasientene ble fulgt opp av en privatpraktiserende lungelege. Disse erfaringene innebærer at dersom man skal gjennomføre CT-screening, må egne ressurser avsettes til arbeidet med oppfølging, både på lungemedisinsk side, og ved røntgenavdelinger. Dette tilsier at man i budsjettet for en CT-screening må beregne inn kostnader ikke bare til selve CT-undersøkelsen, men trolig også til oppfølging innenfor det private helsetilbudet, siden det offentlige systemet neppe har kapasitet.

Vi gjorde også en tredje viktig og alvorlig erfaring. Vi hadde ikke i prosjektet tilgang på en røntgenlege som hadde oppfølging av CT-undersøkelsene som sin primære og dediserte oppgave. For å få utført tyding og klassifisering av filmene var vi avhengige av å samle opp alle filmene og oversende dem til røntgenlegen samlet. På grunn av den tiden selve CT-undersøkelsene tok, og behovet for å frigjøre tid til lesningen, ble de endelige resultatene av tydingen ikke klar før 5-10 måneder etter at den enkelte undersøkelse var gjort. Dette medførte bl.a. at enkelte bifunn, som fortykket aorta (lettgradig aortaaneurisme) ikke ble meddelt deltakerne før flere måneder etter undersøkelsen. Selv om vi ikke har holdepunkter for at denne forsinkelse hadde noen klinisk betydning, skapte det unødig bekymring hos enkelte av deltakerne, som burde vært unngått.

Vi forutsatte videre at ”dramatiske” funn ville bli oppdaget av røntgeninstituttets radiograf ved gjennomgang av filmene etter CT-undersøkelsen, og at det derfor ikke ville være nødvendig med en full beskrivelse av filmen ved røntgenlege ved røntgeninstituttet. Vi ba eksplisitt røntgenavdelingene om å melde fra om det de oppfattet som ”dramatiske” funn, og dette ble som tidligere nevnt gjort i noen få tilfeller, noe som førte til rask oppfølging og utredning. Det viste seg imidlertid at ett tilfelle av lungekreft ble oversett ved denne radiografiske gjennomgangen, og pasienten ble dermed ikke henvist til videre utredning før flere måneder etter at CT-undersøkelsen var gjort. Selv om den utsatte diagnosen ifølge lungelegen trolig ikke har hatt betydning for pasientens prognose, var dette en svært uheldig hendelse, som trolig ville vært unngått ved en ”normal” CT-beskrivelse ved røntgeninstituttets røntgenlege. En slik beskrivelse ville kunnet medføre en tidligere diagnose og oppfølging, også av mindre alvorlige funn, noe som ville ha vært en fordel i hele prosjektet. I ettertid ser vi helt klart at vi burde lagt inn en ordinær CT-beskrivelse av røntgenlege umiddelbart etter undersøkelsen som en del av prosjektet.

I sammenlignende studier mellom røntgen og CT er det godt dokumentert at CT er langt bedre og mer følsom som undersøkelsesmetode^{11;21;22}. Vi kunne derfor kanskje ha

forventninger om at CT kunne gi oss en slags ”sannhet” om lungefunnene hos personene vi undersøkte. I ettertid ser vi at en slik forventning ikke ble innfridd. Erfaringene vi har gjort stemmer godt overens med senere publiserte resultater fra andre studier med bruk av CT, nemlig at CT er en svært sensitiv metode som gir svært mange funn, hvorav en majoritet trolig er falske positive funn. Med falske positive funn mener vi i denne sammenheng funn som ikke har klinisk relevans. Det har vært brukt som et argument for CT-screening at det ved en CT undersøkelse blir gjort flere funn enn ved en røntgenundersøkelse. Dersom over 90 % av disse funnene er falske positive funn, er det grunn til å diskutere om det at det gjøres flere funn er en fordel eller en ulempe med undersøkelsen.

Dette forhold har nylig vært grundig omtalt i en reviewundersøkelse som ble publisert av Midthun et al i 2008²³. Forfatterne adresserer her flere problemer man må ta stilling til i en vurdering av om screening for lungekreft med CT er en aktuell strategi. I 8 studier ble det beskrevet ikke-kalsifiserte noduler hos fra 5-51 % av deltagerne. Andelen av disse nodulene som viste seg å være lungekreft varierte fra 1,7-11,6 %. I den studien med flest falsk positive hadde 36 deltagere lungekreft, mens det ble gjort funn av noduler hos 2099 personer. Prevalensen av lungekreft varierte fra 0,5-2,7 % i disse studiene. I noen prosjekter har man valgt ikke å regne med noduler som er mindre enn 5 mm i diameter, for å redusere antall falsk-positive funn. Dette vil igjen gi en øket risiko for falsk-negative resultater. Funn av falsk-positive noduler vil gi et behov for oppfølgende undersøkelser; i første rekke en kontroll-CT-undersøkelse. I noen av tilfellene vil det også være behov for invasive undersøkelser, biopsi o.l. I Mayo-studien²⁴ ble 18 % av operasjonene (10 av 55) gjort på personer med godartete svulster. I en tysk studie representerte benigne noduler 20 % av reseksjoner som ble utført²⁵. Bach et al.²⁶ fant i en metaanalyse av CT-screening at når antallet kreftsvulster etter screening var 3 ganger det man hadde forventet, var antallet operasjoner 10 ganger det man hadde forventet. Viktige aspekter ved falsk positive funn er kostnader ved ekstra undersøkelser, øket risiko ved operative inngrep på mennesker som i utgangspunktet er friske, og i tillegg et aspekt som ikke er så lett å måle; belastningen på friske personer av å få vite at man har et funn på CT-undersøkelse som kan representere lungekreft. På den annen side vil man ved en regelmessig (årlig) screening kunne legge kontroll av noen av de mindre funnene inn som et ledd i neste års screeningundersøkelse, noe som vil redusere behovet for umiddelbar oppfølging.

Ettersom vi ikke har hatt mulighet i vår undersøkelse til å skaffe endelige tall for hor mange krefttilfeller som er oppdaget, kan vi ikke trekke konklusjoner om relevansen av de funn som er gjort. Vi vet imidlertid at 42 personer er henvist til / anbefalt oppfølging etter funn av noduler. I 15 tilfelle anga minst en CT-leser mistanke om lungekreft, mens det var 5 tilfelle der begge lesere anga lungekreftmistanke. Vi er ikke kjent med hvor mange av disse tilfellene som virkelig hadde lungekreft.

Screening – generelle prinsipper

I diskusjonen om hvorvidt screening skal anbefales bør en refleksjon om prinsipper ved vurdering av screening trekkes inn, og her kan SINTEFs rapport fra 1999 om screening for prostatakreft benyttes²⁷. De oppsummerer Wilson og Jungners²⁸ kriterier i en WHO-rapport fra 1968, som fortsatt regnes som de mest realistiske kriterier:

- Tilstanden skal være vanlig og invalidiserende
- Tilstandens naturlige forløp skal være kjent
- Det skal være en identifiserbar latent eller presymptomatisk fase
- Det skal finnes en pålitelig, valid og reproduserbar screeningstest
- Screeningstesten skal være akseptabel, uten risiko og enkel å utføre
- Screeningstesten skal ha en høy positiv prediktiv verdi
- Screeningstesten skal være sensitiv og spesifikk
- Kostnadene ved screeningprogrammet skal stå i et rimelig forhold til fordelene ved tidlig diagnose
- Det skal finnes en effektiv og tilgjengelig behandling
- Helsevesenet skal ha kapasitet til å behandle tilfellene som identifiseres ved screeningprogrammet
- Det skal være enighet om hvem som skal tilbys behandling

Dersom et screeningprogram ikke oppfyller samtlige av disse kriteriene er det grunn til å stille spørsmål ved om det bør gjennomføres.

Validitet betyr i hvilken grad en screeningstest gir en god preliminær indikasjon på hvilke individer som har og hvilke som ikke har den aktuelle sykdommen. Validitet har to komponenter: Sensitivitet og spesifisitet. *Sensitivitet* uttrykker i hvilken grad målemetoden oppdager personene som har den sykdom/egenskap vi er på jakt etter. Høy sensitivitet er det samme som at det er få falske negative. *Spesifisitet* uttrykker i hvilken grad vi kan identifisere de friske. Høy spesifisitet er det samme som at det er få falske positive.

Det ideelle er at både sensitivitet og spesifisitet er 100 %, men dette er ikke mulig i det praktiske liv. Det kan ikke stilles opp noen generelle nedre grenser for når en metode kan tas i praktisk bruk, men det er vanlig å forlange verdier på over 80 % både for sensitivitet og spesifisitet før en metode kan aksepteres som ledd i masseundersøkelser.

Positiv prediktiv verdi er et mål for hvor mange av dem masseundersøkelsen peker ut som mulig syke som reelt er syke. Dette er ikke bare avhengig av testens sensitivitet og spesifisitet, men også av tilstandens forekomst i befolkningen som blir undersøkt. Hvis sensitiviteten av en metode er lav samtidig som sykdommen er utbredt kan det resultere i at mange pasienter ikke identifiseres. Problematisk er det også når den positive prediktive verdi er lav og sykdommen forekommer sjelden, da dette medfører at få personer i virkeligheten er syke blant de som undersøkelsesmetoden har identifisert som mulig syke.

Potensielle fordeler med screening

- Bedrer prognosen gjennom tidlig diagnostikk og behandling – færre dør av sykdom, redusert sykkelighet ved tidlig behandling
- Trygghetsfølelse for dem som ikke får påvist sykdom gjennom screening og som er sant friske
- Ressursgevinst ettersom kostnadene ved primærbehandling som regel er langt lavere enn kostnadene for en kronisk kreftpasient med sykdom som utvikler seg til terminal fase

Potensielle ulemper ved screening

- For pasienter der prognosen ikke bedres ved tidlig behandling av sykdommen og som derfor lever lengre tid med kreftdiagnosen
- Overbehandling av ”borderline”-diagnoser; dvs tilfeller som ikke vil utvikle seg til en behandlingstrengende tilstand
- Falsk trygghet. Risiko for å neglisjere symptomer under utvikling. Pasienten slår seg til ro med resultatet av screeningundersøkelsen. Resultatet av screening kan derfor for noen bli verre enn uten screening
- Falske positive tester som fører til mer eller mindre omfattende undersøkelser uten at det foreligger noen sykdom
- Undersøkelsesmetoden eller andre diagnostiske undersøkelser av personer som i virkeligheten er friske, vil kunne påføre personene skade
- Kostnadene ved supplerende undersøkelser av tvetydige screeningsvar kan bli høye

Tidlig diagnose har altså flere sider: Tidlig diagnose kan bevirke at pasienten får færre plager som følge av behandling av en mindre avansert sykdom enn om den oppdages på et senere tidspunkt. På den annen side vil pasienter få kunnskap om sin sykdom på et tidligere tidspunkt og dermed leve lenger med vissheten om at han har en mulig dødelig sykdom. Man vet ikke hvilken betydning slik kunnskap har for det enkelte individ og for samfunnet. I forbindelse med at tester og undersøkelser vurderes tatt i bruk i screeningsammenheng, må en derfor gjøre grundige vurderinger av fordeler og ulemper ved en screeningmetode.

Screening – tidlig diagnose av lungekreft

De siste årene er det utkommet en rekke artikler med gjennomgang av fordeler og ulemper ved bruk av CT-undersøkelse ved screening for lungekreft^{23;29;30}. Disse artiklene baserer seg på et omfattende dataunderlag. I den grad det skal tas stilling til spørsmålet om screening med CT-undersøkelse i Norge generelt, og i norsk industri spesielt, er det nødvendig å ta utgangspunkt i denne omfattende litteraturen og de slutninger som er trukket av de store internasjonale undersøkelsene som er gjort.

Konklusjonen så langt, uavhengig av hvilken holdning forfatterne har til screening i utgangspunktet, er entydig: Det er igangsatt omfattende internasjonale studier, hvor flere titalls tusen deltagere skal inkluderes, og hvor hovedformålet med undersøkelsene er å vurdere om screening kan redusere dødelighet av lungekreft, og generell dødelighet. Ingen forfattere anbefaler oppstart av screening for lungekreft, før konklusjonen av disse studiene foreligger.

Det sentrale argumentet for denne tilbakeholdenheten er at selv om pilotstudiene så langt har gitt en økning i antall diagnostiserte lungekrefttilfeller i forhold til forventet, og en økning av levetiden etter lungekreftdiagnose i forhold til forventet, har man ennå ikke lyktes i å vise at den screenete gruppen lever lenger enn dem som ikke blir screenet. Forklaringen på at denne tilsynelatende selvmotsigelsen kan være tilfellet, ligger i begrepene "*lead time bias*", "*length bias*" og "*overdiagnosis*". Disse begrepene blir ikke oversatt til norsk, ettersom de også i norske artikler bare betegnes på engelsk. Mye av fremstillingen i det følgende er basert på en artikkel fra American Cancer Society i 2007 av William C. Black³¹. Denne artikkelen gir en grundig innføring i prinsipper omkring, og oppdatering av eksisterende erfaringer fra lungekreftscreening med bruk av CT.

"*Lead time*" betegner tidsintervallet fra diagnosen blir stilt ved hjelp av en ny diagnostisk metode, til den vil bli stilt ved hjelp av de eldre metodene. Hvis *lead time bias* forklarer en forlenget overlevelsestid etter diagnose, betyr det at sykdommen blir oppdaget tidligere enn den ellers ville ha gjort, og pasienten lever lenger med sin diagnose, men han dør ikke senere enn han ville gjort med de gamle diagnostiske metodene (Figur "2").

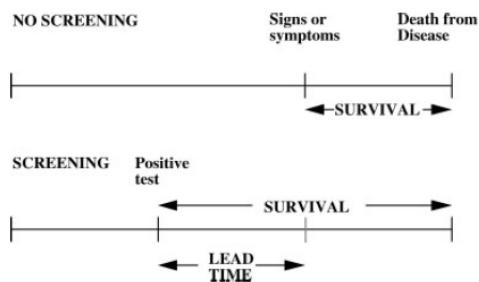


FIGURE 2. Lead time bias. Without screening, diagnosis occurs when clinical signs or symptoms develop. With screening, the time of diagnosis is advanced by lead time provided by a positive test result. If earlier diagnosis has no effect on the time of death from disease, then survival with testing is equal to survival without testing plus lead time. Reprinted with permission from the *American Journal of Roentgenology* (see Black and Welch, 1997²³).

Figuren er hentet fra Black 2007³¹

”Length time bias” handler om at lungekreftsvulster har forskjellig veksthastighet og forskjellig aggressivitet. Når screening foretas i et bestemt tidsintervall, er det større sannsynlighet for å fange de langsomt voksende kreftformene, som kanskje ikke er en like stor trussel for pasienten som de mer aggressive, hurtigvoksende kreftformene. Hurtigvoksende kreftsvulster vil ha større sannsynlighet for enten ikke å ha oppstått, eller å ha kommet over i klinisk fase ved det tidspunkt screeningen foretas (Figur ”3”).

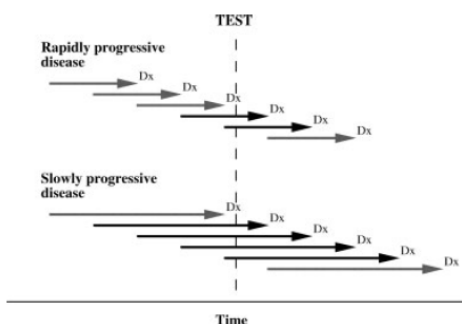


FIGURE 3. Length bias. The probability of detection is related to the rate of disease progression. The length of each arrow represents the length of detectable preclinical phase, from initial detectability to clinical diagnosis (Dx). Testing at a single moment in time detects 4 patients with slowly progressive disease but only 2 patients with rapidly progressive disease (black arrows). Patients who do not have disease detected by the test (gray arrows) are diagnosed clinically, either before or after the time of testing. Reprinted with permission from the *American Journal of Roentgenology* (see Black and Welch, 1997²³).

Figuren er hentet fra Black 2007³¹.

”Overdiagnosis” handler om at screening kan identifisere abnormaliteter som aldri ville gitt pasienten noen klinisk sykdom. Det er to former for overdiagnose: Type I dreier seg om oppdagelse av tidligstadier av sykdom, som enten ikke ville utviklet seg videre, eller som faktisk ville ha gått tilbake av seg selv. Som eksempel på dette nevnes i Black 2007 at en patolog kan feilklassifisere en benign svulst (benign atypisk adenomatøs hyperplasi) som malign (Malignt bronchioalveolært carcinom) – dette er en distinksjon som patologer ofte er uenige om. Type II dreier seg om oppdagelse av tidlige stadier av sykdom som utvikler seg så

langsomt at pasienten ville vært død av andre årsaker før sykdommer hadde gitt kliniske symptomer. Type II overdiagnose forekommer oftest hos individer med langsomt voksende svulster, som har kort forventet levetid på grunn av høy alder eller annen sykdom. I Mayo-studien hadde 13 av 28 screen-oppdagede kreftsvulster en volumetrisk doblingstid på 400 dager. Gitt en gjennomsnittlig diameter på 12 mm, ville en slik kreftsvulst kunne forventes å forårsake død etter mer enn 10 år, og da ville mange av disse personene allerede være døde av andre årsaker, siden den screenede gruppen er personer med høy alder og en betydelig røkehistorie.

I den første Mayo-studien, hvor det til intervensjonsgruppen ble gitt tilbud om screening med røntgenundersøkelse og sputumundersøkelse hver fjerde måned, fant man signifikant økning av antall oppdagede lungekrefttilfeller, økning i kreftsvulster som kunne opereres bort, og forlenget overlevelse etter lungekreftdiagnose i intervensjonsgruppen sammenlignet med kontrollgruppen³². Imidlertid kunne man ikke påvise en signifikant forskjell i lungekreftdødelighet mellom de to gruppene. Marcus et al³³ foretok en utvidet oppfølging av gruppene, med en median oppfølgingstid på 20,5 år, og fant heller ikke her en redusert lungekreftdødelighet i intervensjonsgruppen. Konklusjonen som trekkes i artikkelen er at en del av krefttilfellene som ble avdekket under screening var av begrenset klinisk signifikans.

I en gjennomgang av obduksjonsrapporter fra et sykehus, fant man at 73 % av lungekrefttilfeller som ble identifisert ved obduksjon, var på et tidlig stadium – det var altså ikke denne sykdommen som hadde vært årsak til død hos pasienten³⁴.

På grunn av disse usikkerhetene ved screening, og det faktum at man ikke gjennom de screeningstudiene som hittil er gjennomført har kunnet påvise en reduksjon av dødeligheten i de screenete gruppene, konkluderes det altså med at man så langt ikke kan se bort fra at ulempene ved screening er større enn fordelene. I hvert fall to store studier er igangsatt eller under oppstart, hvor storøykere blir henvist til henholdsvis lavdose CT eller røntgen i den ene studien (USA, 50 000 personer) eller lavdose CT eller vanlig helsetjeneste og råd om røykeslutt i den andre studien (Nederland og Belgia, 16 000 personer). Begge disse studiene har som mål å avdekke om screening gir en redusert dødelighet. Det første studien vil kunne gi resultater fra 2009, den andre fra 2016. Inntil resultatene fra disse studiene foreligger, kan vi ikke på faglig grunnlag tilrå lungekreftscreening ved bruk av CT i yrkeseksponerte høyrisikogrupper.

En av målsettingene med prosjektet var også å vurdere om resultater fra MD-CT-undersøkelse i større grad enn andre metoder gir informasjon som belyser eksponerings-responssammenhenger i SiC-industrien (sammenholde resultater med eksponeringsopplysninger fra jobbeksponeeringsmatrise). Dette forhold er ikke belyst i rapporten, men vil bli nærmere vurdert i forbindelse med videre vitenskapelig bearbeidelse av materialet.

Konklusjoner

CT-prosjektet i silisiumkarbidindustrien

Prosjektet har gitt viktige erfaringer som stemmer godt overens med erfaringer fra større screeningprosjekter internasjonalt. CT-undersøkelsen ga et høyt antall funn, både totalt og for mulige arbeidsrelaterte diagnoser som pneumokoniose, pleuraplakk, emfysem og mulig lungekreft. I alt 30 % av alle deltakerne ble anbefalt videre oppfølging. En stor andel av de aktuelle funn var trolig uten klinisk relevans, og behovet for oppfølging kan ha medført en belastning for den enkelte. Videre var sikkerheten i diagnosen ved CT-undersøkelse alene relativt lav, i og med at samsvaret mellom CT-tyderne, spesielt når det gjelder positiv spesifikt samsvar, var lavt. Dersom et screeningprosjekt med CT-undersøkelse planlegges, må det i tillegg til selve undersøkelsen legges inn planer for og kostnader ved den direkte oppfølgingen av positive funn.

Screening av lungekreft ved hjelp av computertomografi

En rekke internasjonale publikasjoner fremhever usikkerhetene ved CT-screening. Da man i de lungekreftscreeningstudiene som hittil er gjennomført ikke har kunnet påvise en reduksjon av dødeligheten i de screenete gruppene, konkluderes det med at man så langt ikke kan se bort fra at ulempene ved screening er større enn fordelene. Inntil resultater fra store randomiserte studier foreligger, kan vi foreløpig ikke på faglig grunnlag tilrå lungekreftscreening ved bruk av CT i yrkeseksponerte høyrisikogrupper som for eksempel ansatte i silisiumkarbidindustrien.

Screening av andre lungesykdommer ved hjelp av CT

Tidligere studier har konkludert med at CT gir sikrere diagnostikk av støvlungesykdom enn røntgen, men at fordelene ikke er så store at den økede stråledosen og de økede kostnadene ved CT-screening kan forsvares. CT er imidlertid en anbefalt undersøkelse ved individuell oppfølging av usikre røntgenfunn og ved lungesyntomer.

CT-screening i norsk industri

Vi har i studien gjort viktige erfaringer, og vi har i rapporten presentert metodiske svakheter som man dermed senere vil kunne unngå ved en tilsvarende undersøkelse. Imidlertid mener vi at studien gir et realistisk bilde av de utfordringer en vil møte dersom en gjennomføring av lungescreening ved hjelp av MD-CT i regi av f.eks. bedriftshelsetjenester blir innført i norsk industri.

Tillegg

Etter at denne rapporten var ferdigskrevet, ble det avholdt et konsensusmøte mellom de to CT-eksperter. Her ble alle undersøkelser hvor CT-tyderne hadde konkludert forskjellig i første omgang gjennomgått på nytt og diskutert. I forbindelse med møtet ble et nytt ”konsensus-registreringsskjema” utfyllt for 101 personer.

Det er foreløpig ikke utført ytterligere analyser med disse nye dataene. Hovedresultatene fra møtet er likevel verdt å nevne i denne sammenheng: Tyderne oppnådde gjennom dette konsensusmøtet enighet i alle konklusjoner. Dette medførte at ytterligere syv personer ble anbefalt å kontakte sin fastlege for en kontroll-CT-undersøkelse.

Litteratur

- (1) Forskrift om asbest. Arbeidstilsynet, 2005.
<http://www.lovdatab.no/cgi-wift/ldles?doc=/sf/sf/sf-20050426-0362.html>
- (2) Forskrift om vern mot eksponering for kjemikalier på arbeidsplassen. Arbeidstilsynet, 2001.
<http://www.lovdatab.no/cgi-wift/ldles?doc=/sf/sf/sf-20010430-0443.html>
- (3) Andersen A, Høy C. Rapport om sykkelighet av kreft og årsaksspesifikk dødelighet ved silisiumkarbidverk. 77/1988, 1-56. 1988. SIHU dok.
- (4) Romundstad P, Andersen A, Haldorsen T. Cancer incidence among workers in the Norwegian silicon carbide industry. *Am J Epidemiol* 2001;153:978-986.
- (5) Romundstad P, Andersen A, Haldorsen T. Non-malignant mortality among Norwegian silicon carbide smelter workers. *Occup Environ Med* 2002;59:345-347.
- (6) Friedman AC, Fiel SB, Fisher MS, Radecki PD, Lev-Toaff AS, Caroline DF. Asbestos-related pleural disease and asbestosis: a comparison of CT and chest radiography. *Am J Roentgenol* 1988;150:269-275.
- (7) Bergin CJ, Muller NL, Vedal S, Chan-Yeung M. CT in silicosis: correlation with plain films and pulmonary function tests. *Am J Roentgenol* 1986;146:477-483.
- (8) Bergin CJ, Muller NL, Miller RR. CT in the qualitative assessment of emphysema. *J Thorac Imaging* 1986;1:94-103.
- (9) Sanders C, Nath PH, Bailey WC. Detection of emphysema with computed tomography. Correlation with pulmonary function tests and chest radiography. *Invest Radiol* 1988;23:262-266.
- (10) Henschke CI, McCauley D, Yankelevitz DF, Naidich DP, McGuinness G, Miettinen OS et al. Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet* 1999;354:99-105.
- (11) Remy-Jardin M, Degreaf JM, Beuscart R, Voisin C, Remy J. Coal worker's pneumoconiosis: CT assessment in exposed workers and correlation with radiographic findings. *Radiology* 1990;177:363-371.
- (12) Brett GZ. The value of lung cancer detection by six-monthly chest radiographs. *Thorax* 1968;23:414-420.
- (13) Melamed MR, Flehinger BJ, Zaman MB, Heelan RT, Perchick WA, Martini N. Screening for early lung cancer. Results of the Memorial Sloan-Kettering study in New York. *Chest* 1984;86:44-53.
- (14) Lung cancer screening: recommendation statement. *Ann Intern Med* 2004;140:738-739.
- (15) Swensen SJ, Jett JR, Hartman TE, Midthun DE, Sloan JA, Sykes AM et al. Lung cancer screening with CT: Mayo Clinic experience. *Radiology* 2003;226:756-761.
- (16) Hering KG, Tuengerthal S, Kraus T. [Standardized CT/HRCT-classification of the German Federal Republic for work and environmental related thoracic diseases]. *Radiologe* 2004;44:500-511.
- (17) Hripcsak G, Heitjan DF. Measuring agreement in medical informatics reliability studies. *J Biomed Inform* 2002;35:99-110.
- (18) MacMahon H, Austin JH, Gamsu G, Herold CJ, Jett JR, Naidich DP et al. Guidelines for management of small pulmonary nodules detected on CT scans: a statement from the Fleischner Society. *Radiology* 2005;237:395-400.

- (19) Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-174.
- (20) Global Initiative for Chronic Obstructive Lung Disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease (updated 2007). <http://www.goldcopd.com/Guidelineitem.asp?l1=2&l2=1&intId=989>
- (21) Begin R, Ostiguy G, Fillion R, Colman N. Computed tomography scan in the early detection of silicosis. *Am Rev Respir Dis* 1991;144:697-705.
- (22) Gevenois PA, Pichot E, Dargent F, Dedeire S, Vande WR, De VP. Low grade coal worker's pneumoconiosis. Comparison of CT and chest radiography. *Acta Radiol* 1994;35:351-356.
- (23) Midthun DE, Jett JR. Update on screening for lung cancer. *Semin Respir Crit Care Med* 2008;29:233-240.
- (24) Crestanello JA, Allen MS, Jett JR, Cassivi SD, Nichols FC, III, Swensen SJ et al. Thoracic surgical operations in patients enrolled in a computed tomographic screening trial. *J Thorac Cardiovasc Surg* 2004;128:254-259.
- (25) Diederich S, Wormanns D, Semik M, Thomas M, Lenzen H, Roos N et al. Screening for early lung cancer with low-dose spiral CT: prevalence in 817 asymptomatic smokers. *Radiology* 2002;222:773-781.
- (26) Bach PB, Jett JR, Pastorino U, Tockman MS, Swensen SJ, Begg CB. Computed tomography screening and lung cancer outcomes. *JAMA* 2007;297:953-961.
- (27) Fosså SD, Høiseter PÅ, Johansen TB, Vatten L, editors. Screening for prostatakref. Senter for medisinsk metodevurdering, SINTEF. SMM-Rapport 3/1999. ftp://ftp.nokc.no/SMM/SMM-Rapport_99-03_Screening_for_prostatakref..pdf
- (28) Wilson JM, Jungner YG. Principles and practices of screening for disease. 1968. Geneva, WHO.
- (29) Bach PB, Silvestri GA, Hanger M, Jett JR. Screening for lung cancer: ACCP evidence-based clinical practice guidelines (2nd edition). *Chest* 2007;132:69S-77S.
- (30) Smith RA, Cokkinides V, Eyre HJ. American Cancer Society Guidelines for the Early Detection of Cancer, 2005. *CA Cancer J Clin* 2005;55:31-44.
- (31) Black WC. Computed tomography screening for lung cancer: review of screening principles and update on current status. *Cancer* 2007;110:2370-2384.
- (32) Fontana RS, Sanderson DR, Woolner LB, Taylor WF, Miller WE, Muhm JR et al. Screening for lung cancer. A critique of the Mayo Lung Project. *Cancer* 1991;67:1155-1164.
- (33) Marcus PM, Bergstralh EJ, Fagerstrom RM, Williams DE, Fontana R, Taylor WF et al. Lung cancer mortality in the Mayo Lung Project: impact of extended follow-up. *J Natl Cancer Inst* 2000;92:1308-1316.
- (34) McFarlane MJ, Feinstein AR, Wells CK, Chan CK. The 'epidemiologic necropsy'. Unexpected detections, demographic selections, and changing rates of lung cancer. *JAMA* 1987; 258(3):331-338.