

Statistisk analyse av tidligdiagnostikk av kreft krever nye metoder

KRONIKK

PER-HENRIK ZAHL

per-henrik.zahl@fhi.no

Per-Henrik Zahl er cand.med, cand.scient. i matematisk statistikk og dr.med. i biostatistikk.

Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

Metodene for å diagnostisere kreft har tradisjonelt basert seg på at alt vokser. Immunterapi og screeningstudier viser imidlertid at noen svulster forsvinner av seg selv.

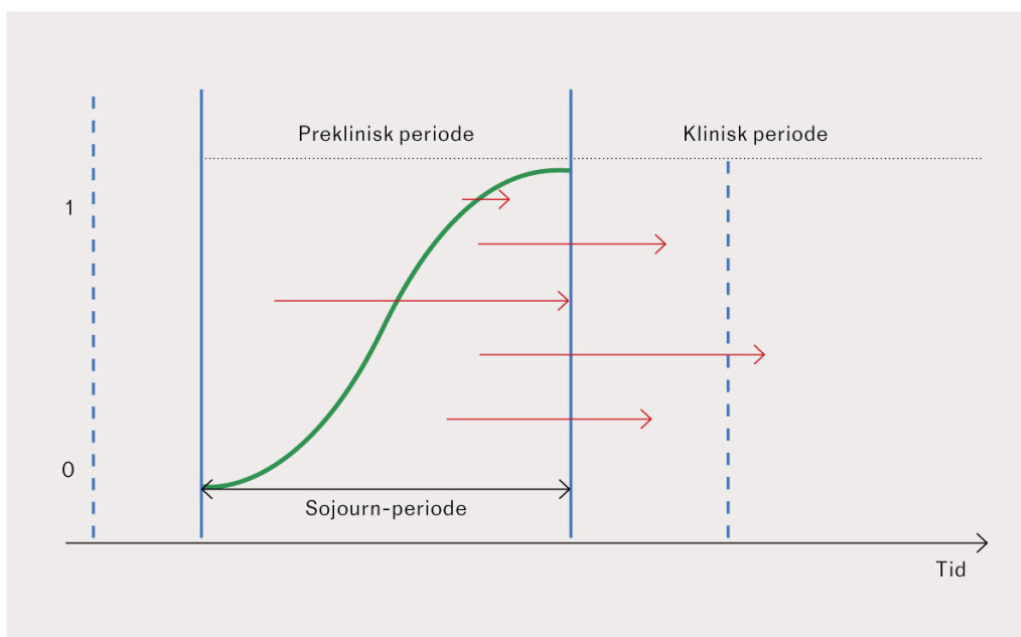
Tradisjonell forståelse av kreft, som danner grunnlaget for diagnostisering, behandling og statistiske analyser, er basert på forestillingen om at svulster vokser kontinuerlig og at spontan regresjon knapt eksisterer. Moderne immunterapi og erfaringer fra masseundersøkelser viser at det ikke alltid er slik. Dette gjør begrepet fremskyndingstid og forestillingen om at tidligdiagnostikk alltid øker sjansen for overlevelse, problematisk.

Fremskyndingstid og preklinisk periode

Estimering av fremskyndingstid (lead time), sensitivitet og preklinisk periode (sojourn time) er sentrale begreper i evaluering av tidligdiagnostikk av kreft. Det finnes forskjellige definisjoner av fremskyndingstid, og det medfører at estimater ikke uten videre kan sammenliknes. Lange fremskyndingstider er ikke statistisk eller biologisk plausible, det er grenser for hvor små svulster man klarer å oppdage. Lange fremskyndingstider gir mye overdiagnostikk, og antagelser om lange fremskyndingstider står sentralt i evaluering av kreftscreening [\(1\)](#).

Fremskyndingstid er definert som tiden diagnosen blir fremskyndet i tid når man innfører en ny type diagnostikk eller når man leter rutinemessig etter en sykdom (2).

Sensitivitet til en undersøkelse er sannsynligheten for at undersøkelsen gjør at man klarer å diagnostisere sykdommen. Da blir (1-sensitivitet) sannsynligheten for ikke å klare å diagnostisere sykdommen, selv om den er til stede, altså sannsynlighet for falskt negative funn (2). Lav sensitivitet kan innebære både lav sannsynlighet for å diagnostisere en sykdom som virkelig er til stede (mange falskt negative), og høy sannsynlighet for at mange små lesjoner man oppdager etter hvert blir borte av seg selv (falskt positive). I det første tilfellet kan undersøkelsen gi falsk trygghet. I det andre tilfellet skapes mye unødvendig angst i tillegg til overbehandling. Det er normalt å anta at sensitiviteten synker når svulstene blir mindre. Jo mindre svulstene er, jo vanskeligere er de å oppdage. Figur 1 illustrerer sammenhengen mellom lengde på fremskyndingstid og sensitivitetsfunksjonen.



Figur 1 Sammenhengen mellom lengde på fremskyndingstid (røde piler) og sensitivitetsfunksjonen (grønn kurve). Sojourn-periode er mellom de to heltrukne blå linjene. Blå stipleddotted linje til venstre er ny diagnostisk metode med lavere deteksjonsterskel, blå stipleddotted linje til høyre er gammel klinisk deteksjonsterskel.

Preklinisk diagnostiserbar periode er definert som intervallet hvor det er teoretisk mulig å stille en diagnose (f.eks. ved innføring av en ny type diagnostikk) før man kan sette diagnosen klinisk. Dette kan også defineres som maksimal fremskyndingstid eller et intervall som omfatter alle mulige fremskyndingstider (fra størst mulig til minst mulig fremskyndingstid) (3). Merk at dette ikke er to identiske definisjoner – den andre definisjonen strekker seg inn i den kliniske perioden. Typisk vil sensitiviteten være minimal eller lik null i starten av den prekliniske perioden, for så å vokse til én på slutten av perioden. Når diagnostiske metoder blir mer sensitive, synker deteksjonsterskelen (illustrert med stipleddotted blå linje til venstre i figur 1).

Ikke all kreft vokser monotont

En matematisk forutsetning for å estimere fremskyndingstid er at sykdommen er monotont voksende (enten eksponentielt, log-normalt eller trinnvis voksende) (2). Svulster kan ikke gå over av seg selv i slike modeller. Noen svulster vokser fort og noen vokser sakte (4). Noen svulster vokser så sakte at individer dør av andre årsaker før sykdommen kan diagnostiseres klinisk eller gir symptomer. Dette kalles overdiagnostikk (5). Det at svulster ikke oppfører seg på en entydig måte, innebærer at tradisjonelle (eksponentielle) vekstmodeller ikke kan beskrive oppførselen til alle svulster. Derfor må hele modellverket revideres. Siden vi ikke på forhånd kan vite hvilke svulster som vil fortsette å vokse og hvilke som vil forsvinne, må vi utvikle en statistisk vekstmodell som åpner for ulike mulige scenarier.

«Siden vi ikke på forhånd kan vite hvilke svulster som vil fortsette å vokse og hvilke som vil forsvinne, må vi utvikle en statistisk vekstmodell som åpner for ulike mulige scenarier»

Estimert fremskyndingstid er et gjennomsnitt. Indolente subkliniske svulster og svulster som går over av seg selv på et subklinisk nivå, har per definisjon ingen fremskyndingstid (de passerer ikke linjen for klinisk deteksjonsterskel), men man kan likevel definere teoretiske fremskyndingstider for slike svulster. Dette vil medføre at gjennomsnittlig fremskyndingstid vokser dramatisk. Slike teoretiske fremskyndingstider gir lite mening, fordi de ikke kan tolkes som tiden diagnosen fremskyndes. Teoretiske fremskyndingstider estimeres ofte for screening og diagnostisering av prostatakreft og brystkreft. Det eneste medisinske fornuftige er å estimere fremskyndingstid for svulster som er progredierende og blir til sykdom. Slike estimater kalles klinisk fremskyndingstid (5).

Man kan enkelt sjekke om det finnes mange indolente svulster eller et reservoar av saktevoksende svulster i autopsistudier (6) eller prevalensstudier ved oppstart av store screeningprogrammer (7).

Test 1 – estimere andel svulster med lang fremskyndingstid

Hvis det finnes mange svulster med lange fremskyndingstider (lengre enn screeningintervallene), så skal man observere en stor prevalenstopp i første screeningrunde (8). Hvis man har screening annethvert år og deteksjonsraten er den samme i første og andre screeningrunde, så er maksimal fremskyndingstid to år. Alle nye svulster oppdaget i andre screeningrunde var ikke diagnostiserbare i første screeningrunde. Hvis man deler raten i andre runde med raten i første runde, får man andelen med fremskyndingstid under to år. Ved innføring av mammografiscreening i Norge og Sverige fant man at

deteksjonsraten i første og andre screeningrunde var omtrent lik for kvinner under 60 år (9, 10). Maksimal fremskyndingstid for disse kvinnene var rundt to år.

Fremskyndingstid og volumdoblingstid

Fremskyndingstid kan lett omregnes til volumdoblingstid. Dermed får man et relativt mål på hvor mye diagnosen fremskyndes i tid i forhold til total veksttid (5). Etter 19 volumdoblinger er diameter rundt 1 mm, men lengde før dette starter svulster å spre seg (11). Rundt 8–9 volumdoblinger senere kan man diagnostisere svulster med de mest sensitive billedteknikkene, slik som MR (tabell 1) (12, 13).

Tabell 1

Antall celledoblinger, antall celler, diameter, volum og medisinsk relevans for brystkreft (11–13).

| Antall celledoblinger | Antall celler | Diameter (mm) | Volum (mm ³) | Medisinsk relevans |
|-----------------------|----------------------|---------------|--------------------------|---|
| 0 | 1 | 0,012 | 0,000001 | Første kreftcelle |
| 19 | 524 288 | 1,0 | 0,52 | Metastasering starter (11) |
| 28 | 268 435 456 | 8,0 | 268 | Mammografisk deteksjonsterskel (13) |
| 29 | 536 870 912 | 10,0 | 536 | Klinisk deteksjonsterskel (13) |
| 31 | 2 147 483 648 | 16,0 | 2 148 | Gjennomsnittlig diameter i RCT-studier (13) |
| 40 | $2,2 \times 10^{12}$ | 135,1 | 1 099 776 | 1 kg og død |

Hvis preklinisk periode ved screening er tre volumdoblinger, vil gjennomsnittlig fremskyndingstid være litt over 1,5 volumdoblinger. Dette setter et tak for hva som er realistiske tider for fremskyndingstid. Hvis volumdoblingstid er ett år, vil det ta rundt 30 år fra start til man kan oppdage svulsten klinisk. Dette er ikke rimelig. Fremskyndingstider på mer enn 1–2 år og prekliniske perioder på 3–7 år er ikke biologisk plausible. Dessuten, normalt er det slik at etter 29 volumdoblinger (når svulsten er klinisk diagnostiserbar) observerer man en avtagende veksthastighet (4). Slike lange fremskyndingstider og avtagende veksthastighet betyr også at folk i snitt vil leve 10–20 år med svulstene etter at de har blitt klinisk diagnostiserbare. Dette er også biologisk urimelig.

Test 2 – lav sensitivitet

Sensitivitet oppgis som et gjennomsnittsestimat for sensitivitetsfunksjonen. Sensitivitet har egentlig to fortolkninger. Som innledningsvis nevnt er dette sannsynligheten for at man klarer å diagnostisere sykdommen. Eller så er (1-sensitivitet) et estimat på tilstander som går over av seg selv. Hvis det finnes et reservoar av svulster med lange fremskyndingstider og som oppdages med lav sensitivitet, så skal dette reservoaret gradvis tømmes over flere screeningrunder og man får en synkende følge av intervallkreftrater (rate av kreft mellom to screeningtester). Hvis intervallkreftraten synker over tid, tyder dette på mange svulster med lang fremskyndingstid som oppdages med lav sensitivitet. Dette er ikke observert for brystkreft.

Test 3 – kompensatorisk fall når screening slutter

Noen svulster vokser så sakte at noen dør av andre årsaker før en screeningoppdaget sykdom blir til klinisk sykdom. Antallet beregnes ved å regne ut kumulativ insidens for en fødselskohort fra start av screening til rundt ti år etter at screeningen er avsluttet. Nesten all økning i forekomst av sykdom under en screeningperiode skal kompenseres av en reduksjon i forekomst når man ikke lenger går til screening. Ved mammografiscreening skal det normalt bare bli rundt 2 % mer brystkreft når fremskyndingstiden er 4,8 år (8). Etter en stor økning i forekomst ved screening skal man etterpå observere et like stort kompensatorisk fall i forekomsten (8). Observert forekomst i ti år etter at screeningen er avsluttet, sammenliknes med forventet forekomst uten screening.

Justering for fremskyndingstid

Insidensrater og overdiagnostikk skal bare justeres for klinisk fremskyndingstid (5). Det er viktig å skille mellom voksende forekomst av klinisk kreft og voksende forekomst av overdiagnostiserte svulster. Det er ingen grunn til å tro at risikofaktorene er de samme for overdiagnostiserte svulster som for klinisk kreft. Overdiagnostikk kan derfor være en viktig konfunderende faktor og forklare hvorfor mange risikofaktorer for en kreftdiagnose ikke er risikofaktorer for å dø av kreftdiagnosen (14).

«Det er viktig å skille mellom voksende forekomst av klinisk kreft og voksende forekomst av overdiagnostiserte svulster»

Hvis individer har mange indolente svulster i prostata (15) og i nyrene (16) eller svulster i nyrene (16) og nevroblastom (17) som går over av seg selv, så gir screening opplagt mye overdiagnostikk. Spontan regresjon er mye mer vanlig

enn hva leger flest tror – dette kan man også konkludere ut fra effekten av moderne immunterapi, som er basert på dette prinsippet (18).

Justering for fremskyndingstid er en tvilsom statistisk praksis, fordi estimert fremskyndingstid varierer med hvordan fremskyndingstid er definert, med diagnostisk metode og hvor mye leger og individer leter etter små kliniske svulster. Dessuten, fremskyndingstid er spekulativt fordi det er en ikke-observerbar størrelse. Det er ingen grunn til å tro at justering for fremskyndingstid reduserer bias, det kan like godt øke bias.

REFERENCES

1. The Research Council of Norway. Research-based evaluation of the Norwegian Breast Cancer Screening Program.
<https://www.forskningsradet.no/siteassets/publikasjoner/1254012138940.pdf> Lest 30.4.2024.
2. Day NE, Walter SD. Simplified models of screening for chronic disease: estimation procedures from mass screening programmes. *Biometrics* 1984; 40: 1–14. [PubMed][CrossRef]
3. Oxford reference. Sojourn time.
<https://www.oxfordreference.com/display/10.1093/oi/authority.20110803100516669> Lest 12.3.2024.
4. Spratt JA, von Fournier D, Spratt JS et al. Decelerating growth and human breast cancer. *Cancer* 1993; 71: 2013–9. [PubMed][CrossRef]
5. Zahl P-H, Jørgensen KJ, Gøtzsche PC. Overestimated lead times in cancer screening has led to substantial underestimation of overdiagnosis. *Br J Cancer* 2013; 109: 2014–9. [PubMed][CrossRef]
6. Nielsen M, Jensen J, Andersen J. Precancerous and cancerous breast lesions during lifetime and at autopsy. A study of 83 women. *Cancer* 1984; 54: 612–5. [PubMed][CrossRef]
7. Smith-Bindman R, Chu PW, Miglioretti DL et al. Comparison of screening mammography in the United States and the United kingdom. *JAMA* 2003; 290: 2129–37. [PubMed][CrossRef]
8. Boer R, Warmerdam P, de Koning H et al. Extra incidence caused by mammographic screening. *Lancet* 1994; 343: 979 [Letter].. [PubMed][CrossRef]
9. Zahl P-H, Maehlen J, Welch HG. The natural history of invasive breast cancers detected by screening mammography. *Arch Intern Med* 2008; 168: 2311–6. [PubMed][CrossRef]
10. Zahl P-H, Gøtzsche PC, Mæhlen J. Natural history of breast cancers detected in the Swedish mammography screening programme: a cohort study. *Lancet Oncol* 2011; 12: 1118–24. [PubMed][CrossRef]

11. Folkman J, Watson K, Ingber D et al. Induction of angiogenesis during the transition from hyperplasia to neoplasia. *Nature* 1989; 339: 58–61. [PubMed][CrossRef]
 12. Bakker MF, de Lange SV, Pijnappel RM et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. *N Engl J Med* 2019; 381: 2091–102. [PubMed][CrossRef]
 13. Gøtzsche PC, Jørgensen KJ, Zahl P-H et al. Why mammography screening has not lived up to expectations from the randomised trials. *Cancer Causes Control* 2012; 23: 15–21. [PubMed][CrossRef]
 14. Strand BH, Tverdal A, Claussen B et al. Is birth history the key to highly educated women's higher breast cancer mortality? A follow-up study of 500,000 women aged 35-54. *Int J Cancer* 2005; 117: 1002–6. [PubMed][CrossRef]
 15. Draisma G, Boer R, Otto SJ et al. Lead times and overdetetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. *J Natl Cancer Inst* 2003; 95: 868–78. [PubMed][CrossRef]
 16. Jewett MA, Mattar K, Basiuk J et al. Active surveillance of small renal masses: progression patterns of early stage kidney cancer. *Eur Urol* 2011; 60: 39–44. [PubMed][CrossRef]
 17. Schilling FH, Spix C, Berthold F et al. Neuroblastoma screening at one year of age. *N Engl J Med* 2002; 346: 1047–53. [PubMed][CrossRef]
 18. The nobel prize. The Nobel Prize in Physiology or Medicine 2018. <https://www.nobelprize.org/prizes/medicine/2018/press-release/> Lest 12.3.2024.
-

Publisert: 28. mai 2024. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.24.0153
Mottatt 14.3.2024, første revisjon innsendt 18.4.2024, godkjent 30.4.2024.
Opphavsrett: © Tidsskriftet 2026 Lastet ned fra tidsskriftet.no 24. juni 2026.