
Troværdigheden af radiologisk cariesdiagnostik – sammenhæng mellem rigtighedsværdi og sandhedsreference

HANNE HINTZE

En stor del af den odontologiske radiologiske forskning er koncentreret omkring evaluering af nye røntgenmodaliteter anvendt til diagnostik af forskellige sygdomme. I de forgangne år er der således gennemført adskillige undersøgelser for evaluering af den cariesdiagnostiske formåen af de mange nye digitale røntgensystemer. Undersøgelserne har blandt andet til formål at tilvejebringe information, der kan bruges af klinikerens, når beslutning om eventuel udskiftning af det traditionelle filmbaserede røntgensystem med et digitalt skal tages.

Cariesdiagnostik

Idealmålet for en cariesdiagnostisk metode er, at den er i stand til at identificere 'sandt cariøse' og 'sandt ikke-cariøse' tandflader. Med udtrykkene 'sandt cariøse' henholdsvis 'sandt ikke-cariøse' flader menes flader, som i virkeligheden har caries henholdsvis i virkeligheden er sunde. For at afgøre dette må tandfladerne undersøges med en metode, som er i stand til at vise sandheden. Ofte betegnes en sådan metode for sandhedsreferencen, valideringsmetoden eller det diagnostiske facit og på engelsk 'the gold

standard'. Med det diagnostiske facit som reference kan den diagnostiske formåen (rigtigheden) af en metode, som ønskes bedømt, derefter beregnes.

Sandhedsreferencer

En stor del af de videnskabelige undersøgelser, der søger at bedømme rigtigheden af forskellige røntgensystemer anvendt til diagnostik af caries, gennemføres som laboratorieundersøgelser på humane ekstraherede tænder. På sådanne tænder er det muligt ved vejs ende at foretage en gennemskæring for at bedømme deres snitflader histologisk for tilstedeværelse af caries og dermed fremsætte facit. En histologisk sandhedsreference opfylder stort set de tre krav, der gælder for enhver valideringsmetode^{1,2}:

1. Den sande diagnose skal stilles ved hjælp af de kriterier, der definerer sygdommen – når der er tale om caries, skal valideringsmetoden således være i stand til at påvise demineralisering.
2. Den sande diagnose skal være reproducerbar.
3. Den sande diagnose skal være uafhængig af den metode, rigtigheden bedømmes af – er det rigtigheden af en røntgenmetode, som ønskes bedømt, må valideringsmetoden således ikke også være en røntgenmetode.

En stor ulempe ved den histologiske valideringsmetode er imidlertid, at den er destruktiv og dermed kun kan anvendes på ekstraherede tænder eller tænder *in situ*, som er planlagt til efterfølgende ekstraktion. Nu til dags omfatter ekstraherede tænder i den vestlige verden stort set kun 3. molarer. Hvorvidt sådanne molarer er repræsentative også for de andre tandtyper i munden er dog tvivlsomt. Dermed kan der være problemer med klinisk generalisering til hele tandsættet af resultater fra laboratoriestudier, der bedømmer forskellige metoders cariesdiagnostiske formåen fortrinsvis på 3. molarer. Fordelen ved at anvende

ekstraherede humane tænder er imidlertid, at de carieslæsioner, som tænderne besidder, er udviklet biologisk og dermed fremstår, som carieslæsioner gør i virkeligheden.

Studiedesigns, hvor der med bor skabes defekter i ekstraherede tænders approssimalflader som substitut for biologisk udviklede carieslæsioner, har i perioder været anvendt i videnskabelige undersøgelser af røntgensystemers cariesdiagnostiske rigtighed². Med et sådant design er det let at etablere facit, men det er svært at hævde, at sådanne boredefekter ligner biologisk udviklede carieslæsioner. Dermed er det problematisk at overføre resultaterne til klinikken, hvor biologisk udviklede carieslæsioner ikke fremstår som defekter skabt med en boremaskine³. Enkelte undersøgelser har da også vist, at rigtighedsværdierne for identifikation af boredefekter er væsentligt højere end dem, som findes for diagnostik af biologisk udviklede carieslæsioner^{4,5}. Dermed synes det at være klart, at studiedesigns med boredefekter som substitut for 'rigtige' carieslæsioner ikke tilvejebringer ny klinisk relevant information.

For at resultater fra videnskabelige undersøgelser kan overføres til klinikken, er det vigtigt, at undersøgelserne gennemføres under forhold, der ligner virkeligheden mest muligt. Det indebærer, at de metoder, som benyttes, anvendes under *in vivo*-betingelser på tænder *in situ* med biologisk udviklede carieslæsioner. Problemet med et sådant design er imidlertid, at det er yderst vanskeligt at konstatere det diagnostiske facit vedrørende cariestilstedeværelsen, da det er umuligt efterfølgende at ekstrahere de undersøgte tænder for at gennemføre histologisk validering. Sidstnævnte problem har nogle forskere forsøgt at løse ved at lade eksperter tolkning af røntgenbilleder af de undersøgte tænder tjene som facit². At anvende røntgenbilleder for at fastlægge den sande cariestilstedeværelse er dog ikke hensigtsmæssigt, da det på sådanne ikke er muligt at identificere mere end godt halvdelen af de læsioner, der reelt findes (bedømt histologisk). I en undersøgelse bedømte to røntgeneksperter tilstedeværelsen af caries på 373 approssimalflader ud fra konventionelle røntgen-

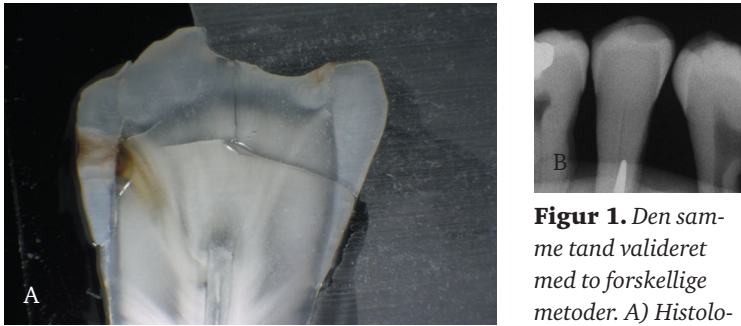
Tabel 1. Histologisk og radiologisk cariesvalidering af 373 approximalflader

	Sund n	Emaljecaries n	Dentincaries n	Total n
Histologisk validering	208	113	52	373
Radiologisk validering	278	42	53	373

Resultaterne hidrører fra en undersøgelse af Hintze & Wenzel⁶

billeder og ud fra en histologisk undersøgelse, efter at fladerne var opskåret i tynde snit⁶. Resultatet blev, at der på basis af den histologiske undersøgelse fandtes 165 carieslæsioner, hvoraf 52 var lokaliseret til dentinen (Tabel 1). Af det totale antal carieslæsioner bedømtes 48% tilsvarende ud fra røntgenbillederne, mens det for dentinlæsionernes vedkommende var 53%.

At der er forskel på en histologisk og en radiologisk cariesvalidering fremgår også af Figur 1, hvor den samme tand ses med de to undersøgelsesmetoder. På venstre approximalflade ses ingen caries med den radiologiske metode – men tydelig dentincaries med den histologiske. Skulle der derimod have været en radiolucet forandring på røntgenbilledet, ville det imidlertid ikke være et sikkert tegn på caries, da der radiologisk ikke kan skelnes éntydigt imellem caries og hypomineralisering, præparation, plastfyldning og lignende. En anden alvorlig fejl ved at bruge eksperter vurdering af røntgenbilleder som sandhedsreference for røntgensystemer er, at sandhedsreference og metode under bedømmelse ikke er uafhængige, som forskrifterne kræver, at de skal være. At bedømme en metode og samtidig anvende den som facit er i følge Freedman⁷ det samme som at tillade en eksamenskandidat at rette en del af sine egne eksamensopgaver. Noget sådant vil ingen tillade af frygt for rettebias.



Figur 1. Den samme tand valideret med to forskellige metoder. A) Histologisk facit visende dentincaries på venstre approximalflade. B) Radiologisk facit (tanden i centrum) visende fravær af caries (sund) på venstre approximalflade.

Terminologi for diagnostisk rigtighed

Som udtryk for en diagnostisk metodes rigtighed benyttes ofte simple talværdier, som fortæller, hvor effektiv metoden er til at identificere syge henholdsvis raske individer/tandflader. Når det drejer sig om metoder, hvis udbytte baserer sig på en human observatørfortolkning, er metodens samlede rigtighed en blanding af metodens effektivitet og observatørens evne til at drage korrekte tolkninger.

Sensitivitet og specificitet

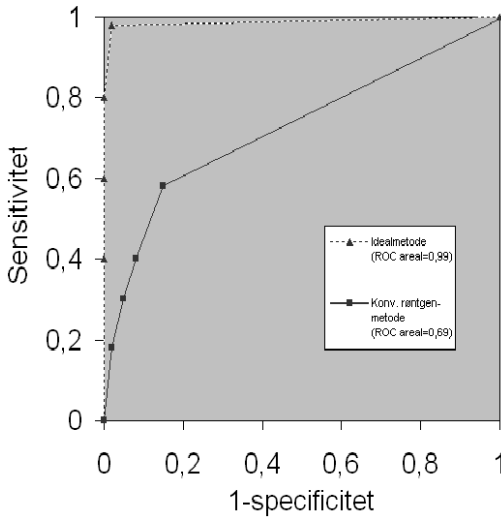
Traditionelt anvendes parametrene sensitivitet og specificitet som udtryk for en metodes rigtighed. I forbindelse med cariesdiagnostik udtrykker sensitiviteten sandsynligheden for en positiv diagnose på en cariøs tandflade. Specificiteten udtrykker sandsynligheden for en negativ diagnose på en sund tandflade. En fejlfri metode har en sensitivitet på 1,00 og en specificitet på 1,00 som udtryk for, at metoden kun giver anledning til sande diagnoser (cariøse flader identificeres som cariøse, og sunde flader identificeres som ikke-cariøse). I virkelighedens verden findes der imidlertid ingen fejlfrie cariesdiagnostiske undersøgelsesmetoder. Det er dog enhver klinikers ønske at benytte metoder

med størst mulig træfsikkerhed, det vil sige metoder med både en høj sensitivitet og en høj specificitet (gerne 0,80 og derover). At opnå en perfekt sensitivitet eller en perfekt specificitet er dog ingen kunst. Det kræver blot, at der registreres caries eller fravær af caries på samtlige tandflader, der undersøges. Men at opnå en høj værdi for begge parametre er en udfordring, og for mange metoder ses, at hvis sensitiviteten er høj, så er specificiteten lav eller omvendt. Kendes værdierne for forskellige metoder, kan klinikerne imidlertid vælge den metode, der har de sensitivitets- og specificitetsværdier, som bedst opfylder det ønskede.

ROC areal

En anden ofte anvendt parameter for beskrivelse af en metodes rigtighed i forhold til sandhedsreferencen er ROC (Receiver Operating Characteristic) arealet. Det er arealet under en kurve i et to-akse diagram, hvor sensitiviteten afsættes op ad Y-aksen og specificiteten ud ad X-aksen, almindeligvis som 1-specificiteten. For en række forskellige diagnostiske tærskler (helt sikkert caries, næsten sikkert caries, og så videre) kan sensitivitets- og 1-specificitetsværdierne for en given metode plottes ind i ROC diagrammet. Ved forbindelse af plot-punkterne fremkommer en såkaldt ROC kurve, som viser sandsynligheden for positiv cariesdiagnostik på cariøse flader i kombination med sandsynligheden for diagnostik af caries på sunde flader (indirekte sandsynligheden for negativ cariesdiagnostik på sunde flader) ved forskellige diagnostiske tærskler. Ideelt set bør en ROC kurve smygge sig op langs Y-aksen, bøje af i øverste venstre hjørne og forløbe langs toppen af diagrammet og ende i øverste højre hjørne (Figur 2). Dermed vil arealet under kurven, der bruges som udtryk for metodens overordnede diagnostiske rigtighed, blive 1,00, som er værdien for den fejlfri metode.

En fordel ved anvendelse af ROC arealet som udtryk for en metodes rigtighed er, at der så blot er én talværdi at forholde sig til i modsætning til ved brug af sensitivitet og specificitet, hvor der er to talværdier at tage hensyn til på en og samme tid.



Figur 2. ROC kurver for en ideal diagnostisk metode og en typisk røntgenmetode.

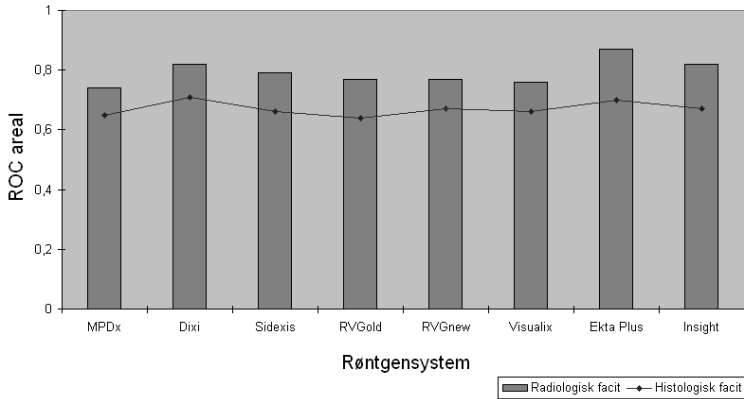
Hvad kan influere på rigtighedsværdien af en radiologisk metode anvendt til cariesdiagnostik?

Fra laboratorieundersøgelser, der har anvendt histologisk validering som facit for tilstedeværelse af carieslæsioner, findes ofte sensitivitets- og specificitetsværdier, som er langt under 1,00. White & Yoon⁸ fandt således sensitivitetsværdier mellem 0,50 og 0,60 og specificitetsværdier omkring 0,80 for diagnostik af approksimale carieslæsioner med et konventionelt og et digitalt røntgensystem. Wenzel *et al.*⁹ fandt sensitivitetsværdier mellem 0,62 og 0,72 og specificitetsværdier mellem 0,76 og 0,90 for diagnostik af dentin carieslæsioner på okklusalflder med et konventionelt og flere digitale røntgensystemer. Udtrykt ved hjælp af ROC arealet synes mange af de tilsvarende røntgensystemer at opnå værdier mellem 0,55-0,88 for diagnostik af approksimalfledarier og 0,76-0,82 for diagnostik af okklusalfledarier². Samlet fortæller rigtighedsværdier af sådanne størrelser, at røntgen som diagnostisk værktøj er en middelmådig metode. Den er på ingen måde fejlfri, idet den giver anledning til mange falske

negative registreringer (læsioner, som findes i virkeligheden, men som ikke opdages) samtidig med, at den resulterer i falske positive diagnoser (flader, som i virkeligheden er sunde, bedømmes til at være cariøse).

Meget høje rigtighedsværdier, især ROC arealer, for røntgensystemer anvendt til diagnostik af caries er imidlertid rapporteret fra enkelte studier. Der findes således eksempler på, at der på digitale røntgenbilleder tilført støj og destruktiv kompression er opnået ROC arealer på 0,92-0,96¹⁰. Umiddelbart indikerer så høje værdier, at billederne har en næsten perfekt diagnostisk rigtighed, idet forskellen op til det maksimalt opnåelige er ubetydelig. Læsere af sådanne resultater kan således forledes til at tro, at anvendelsen af digitale røntgenbilleder tilført støj og destruktiv kompression giver anledning til en signifikant bedre diagnostisk rigtighed end anvendelse af konventionelle røntgenbilleder, som almindeligvis giver anledning til meget mindre ROC arealer, selvom dette strider mod al fornuftig logik. Den direkte årsag til de meget høje ROC arealer er da heller ikke bedre rigtighed, men blot at sandhedsreferencen for tilstedeværelse af caries ikke har været histologisk undersøgelse. Derimod har en radiologisk validering, hvor eksperter på baggrund af konventionelle røntgenbilleder har stillet den „sande“ diagnose, været anvendt, og derefter har disse eksperters validering været anvendt som facit for en gruppe tandlægers bedømmelse af carieslæsioner på de tilsvarende digitale røntgenbilleder tilføjet varierende mængder billedødelæggelse. I videnskabelige undersøgelser, der anvender identiske metoder for bedømmelse og facit, synes der således at være stor sandsynlighed for at opnå meget høje rigtighedsværdier – værdier, der kan nærme sig maksimumværdien.

Denne påstand understøttes af et studie, hvor rigtigheden af otte røntgensystemer bedømtes¹¹. I dette studie blev der anvendt to forskellige sandhedsreferencer – en histologisk og en radiologisk. Den radiologiske sandhedsreference gav anledning til signifikant højere ROC arealer for samtlige otte systemer under bedømmelse, end når den histologiske sandhedsreference anvend-



Figur 3. Diagnostisk rigtighed for opdagelse af approksimalladecaries udtrykt ved hjælp af ROC arealet for otte røntgensystemer. Arealerne for hvert system er beregnet i relation til en histologisk og en radiologisk facitliste (resultaterne hidrører fra en undersøgelse af Hintze & Wenzel¹¹).

tes (Figur 3). I gennemsnit var ROC arealerne for diagnostik af approksimalladecaries beregnet på basis af den radiologiske sandhedsreference 18% højere end dem, som beregnedes på basis af den histologiske. Ud over denne forskel viste der sig også at være stor forskel på den indbyrdes rangorden af de otte røntgensystemer alt efter, hvilken metode der anvendtes som facit².

Kvaliteten af den radiologiske facitliste har også vist sig at være af betydning for det ROC areal, som en metode under bedømmelse opnår. Wenzel & Hintze¹² viste således, at ROC arealet for digitale fosforplade-baserede røntgenbilleder, anvendt til diagnostik af approksimalladecaries, var 0,66, når den sande cari-estilstedeværelse konstateredes ved brug af en histologisk validering. Anvendtes derimod en radiologisk facitliste baseret på røntgenregistreringer foretaget af den dårligste af fem røntgeneksperter, blev det tilsvarende ROC areal beregnet til 0,91 – 38% højere. Anvendtes i stedet røntgenregistreringer foretaget af den bedste af fem røntgeneksperter, beregnedes ROC arealet til 0,74 – 12% højere end det, der fremkom på basis af den histologiske

facitliste. Der synes dermed at være en tendens til højere ROC areal desto ringere diagnostiker, der anvendes for at konstatere sandheden.

Konklusion og konsekvens

Konklusionen på ovennævnte er, at sandhedsreferencen for cariesvurdering har stor indflydelse på rigtighedsværdierne af radiologiske systemers diagnostiske formåen. Jo mindre robust sandhedsreference og eventuelt dårligere diagnostiker, desto højere rigtighedsværdier, og ofte bliver rigtighedsværdierne så høje, at de nærmer sig det perfekte.

Konsekvensen af denne sammenhæng er, at læsere af videnskabelige artikler og salgsmaterialer vedrørende metoders diagnostiske formåen skal være på vagt over for meget høje rigtighedsværdier. Man bør således altid få oplysning om, hvilken metode der har været anvendt til cariesverificeringen. Er det ikke en metode, som opfylder de tre universelle krav til en robust valideringsmetode^{1,2}, bør man være yderst skeptisk over for de rigtighedsværdier, man præsenteres for.

LITTERATUR

- 1 Gjørup T. Klinisk vurdering af diagnostiske undersøgelsesmetoder. (Disputats) Lægeforeningens Forlag, 1988.
- 2 Hintze H. Radiography for the detection of dental caries lesions. (Disputats) University of Aarhus, 2004.
- 3 Farman AG. Standards in intraoral radiographic imaging. *Dentomaxillofac Radiol* 2000;29:257-259.
- 4 Price C, Ergül N. A comparison of a film-based and a direct digital dental radiographic system using a proximal caries model. *Dentomaxillofac Radiol* 1997;26:45-52.
- 5 Kang B-C, Goldsmith LJ, Farman AG. Observer differentiation of mechanical defects versus natural dental caries cavitations on mon-

- itor-displayed images with imaging plate readout. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1998;86:595-600.
- 6 Hintze H, Wenzel A. Diagnostic outcome of methods frequently used for caries validation. A comparison of clinical examination, radiography and histology following hemisectioning and serial tooth sectioning. *Caries Res* 2003;37:115-124.
 - 7 Freedman LS. Statistical methods of evaluating and comparing imaging techniques. In: Bleeheh NM. *Investigational techniques in oncology*. Springer-Verlag, Berlin, 1987, s 35-53.
 - 8 White SC, Yoon DC. Comparative performance of digital and conventional images for detecting proximal surface caries. *Dentomaxillofac Radiol* 1997;26:32-38.
 - 9 Wenzel A, Hintze H, Mikkelsen L, Mouyen F. Radiographic detection of occlusal caries in noncavitated teeth. A comparison of conventional film radiographs, digitized film radiographs, and RadioVioGraphy. *Oral Surg Oral Med Oral Pathol* 1991;72:621-626.
 - 10 Janhom A, van der Stelt PF, van Ginkel FC, Geraets WGM. Effect of noise on the compressibility and diagnostic accuracy of caries detection on digital bitewing radiographs. *Dentomaxillofac Radiol* 1999;28:6-12.
 - 11 Hintze H, Wenzel A. Influence of the validation method on diagnostic accuracy for caries. A comparison of six digital and two conventional radiographic systems. *Dentomaxillofac Radiol* 2002;31:44-49.
 - 12 Wenzel A, Hintze H. Comparison of microscopy and radiography as gold standards in radiographic caries diagnosis. *Dentomaxillofac Radiol* 1999;28:182-185.
