



# **Inventarisatie leesstrategieën BE's**

Discussiestuk voor Werkgroep Kwaliteit, 16 mei 2017

Mei 2017

T.D. Geertse

Prof Dr. R.M. Pijnappel

## Inleiding

Op 23 oktober 2015 is Lisa Klompenhouwer aan de Universiteit van Maastricht gepromoveerd op het onderzoek “Strategies for the improvement of population based breast cancer screening.” In haar proefschrift zijn verschillende mogelijke optimalisaties van het huidige beoordelingsproces onderzocht. Het Centrum voor Bevolkingsonderzoek (CvB) heeft het Landelijk Referentiecentrum voor Bevolkingsonderzoek (LRCB) in 2016 gevraagd een notitie te schrijven met daarin een duiding van de onderzoeksresultaten.

Deze gevraagde notitie is op 14 november 2016 aangeboden aan het CvB. De notitie eindigt met een lijst van voorstellen en overwegingen vanuit het LRCB. Als vervolg op deze notitie heeft het CvB, eind 2016, aan het LRCB gevraagd om onder de 15 beoordelingseenheden (BE's) binnen de borstkankerscreening een vragenlijst uit te zetten om de huidige leesstrategie van deze BE's te inventariseren.

Dit document geeft een overzicht van de uitkomsten van deze inventarisatie. Tevens beschrijft het document welke leesstrategie op dit moment volgens het LRCB de voorkeur heeft en welke punten volgens het LRCB verder belicht zouden moeten worden.

## Methode: De vragenlijst

Aan de coördinerend radiologen van alle 15 BE's is een vragenlijst opgestuurd met de volgende vragen:

1. Wordt er binnen de BE (deels) gescreend in vaste koppels?
2. Zo ja, hoeveel koppels op het totaal?
3. Zo ja, worden (individuele) resultaten van verwijscijfers en detectiecijfers gebruikt om deze koppels samen te stellen?
4. Zo nee, graag beschrijven welke redenen er zijn om niet in vaste koppels te screenen:
5. Wordt er bij het inplannen rekening gehouden met de verdeling 1e en 2e lezingen van een screeningsradioloog?
6. Hoeveel dagproducties worden er (gemiddeld genomen) per sessie gelezen door een screeningsradioloog?
7. Wordt er volledig blind gelezen, of wordt er op een of andere manier informatie aan elkaar doorgegeven? Zo ja, dan graag beschrijven op welke wijze:
8. Hoe wordt er om gegaan met discrepanties? Graag de procedure beschrijven:
9. Welk hanging protol wordt er gebruikt?
  - a. Het door de fabrikant ingestelde hanging protocol
  - b. Het door de screeningsorganisatie ingestelde hanging protocol
  - c. Het door de groep/maatschap ingestelde hanging protocol
  - d. Alle screeningsradiologen van de groep hebben een persoonlijk ingesteld hanging protocol
  - e. Onbekend
10. Wordt de toggle/blinken functie door de screeningsradiologen gebruikt?
  - a. Door iedereen
  - b. Door het grootste gedeelte van de groep
  - c. Door een enkeling van de groep
  - d. Door niemand van de groep
  - e. Onbekend

Van de 15 BE's hebben 13 BE's de vragenlijst ingevuld geretourneerd. Twee coördinerend radiologen (van BE Groningen en BE Rotterdam) hebben ook na herhaaldelijke verzoeken (via email en telefonisch) niet gereageerd.

In de volgende hoofdstukken geven tabellen een overzicht van de resultaten. De vragen zijn per onderwerp bij elkaar genomen: het screenen in koppels, de omvang van de sessie, double reading, omgaan met discrepanties en het hanging protocol. Om uit de resultaten betere conclusies te kunnen trekken is in de tabellen "omvang sessie", "double reading" en "omgaan met discrepanties" informatie uit LRCB visitaties toegevoegd (meest recente visitatie). In de tabellen zijn de BE's gecodeerd.

## Resultaten: Screenen in koppels

	Vraag 1	Vraag 2	Vraag 3	Vraag 4
BE	<b>Wordt er binnen de BE (deels) gescreend in vaste koppels?</b>	<b>Zo ja, hoeveel koppels op het totaal?</b>	<b>Zo ja, worden (individuele) resultaten van verwijscijfers en detectiecijfers gebruikt om deze koppels samen te stellen?</b>	<b>Zo nee, graag beschrijven welke redenen er zijn om niet in vaste koppels te screenen:</b>
B	Nee	n.v.t.	n.v.t.	<ul style="list-style-type: none"> <li>• Selectiecriteria voor vorming koppels nochtans onvoldoende inzichtelijk.</li> <li>• Praktisch lastig uitvoerbaar</li> </ul>
C	Eigenlijk niet. Voorheen namen 1 of 2 radiologen een fors deel van het werk voor hun rekening namen. Daardoor ontstonden er koppels.	n.v.t.	Nee dus, maar dat berust niet in principiële onwil.	<ul style="list-style-type: none"> <li>• Onduidelijkheid over de meerwaarde van deze vaste koppels.</li> <li>• Praktisch feitelijk onmogelijk.</li> </ul>
D	Nee	n.v.t.	n.v.t.	Organisatorisch. Heeft geen voorkeur
E	Nee	n.v.t.	n.v.t.	Onze detectiecijfers lopen weinig uiteen.
F	Nee, in principe niet. Echter momenteel in een overgangsfase waarbij 4 van de 8 radiologen afscheid gaan nemen/ hebben genomen. Wens is deze overgang geleidelijk te laten plaatsvinden. Enkele oudere radiologen blijven na hun pensioen nog 1 of 2 jaar doorscreenen en worden dan gekoppeld aan nieuwe (onervaren) radiologen.	n.v.t.	n.v.t.	Praktisch niet mogelijk.

	Vraag 1	Vraag 2	Vraag 3	Vraag 4
BE	<b>Wordt er binnen de BE (deels) gescreend in vaste koppels?</b>	<b>Zo ja, hoeveel koppels op het totaal?</b>	<b>Zo ja, worden (individuele) resultaten van verwijscijfers en detectiecijfers gebruikt om deze koppels samen te stellen?</b>	<b>Zo nee, graag beschrijven welke redenen er zijn om niet in vaste koppels te screenen:</b>
G	Nee	n.v.t.	n.v.t.	<ul style="list-style-type: none"> <li>• Planningstechnisch.</li> <li>• Vermijden op elkaar “ingespeeld raken” van de radiologen in een vast koppel.</li> <li>• Onervaren radiologen strategieën van diverse ervaren radiologen te laten ervaren.</li> </ul>
H	Nee	n.v.t.	n.v.t.	<ul style="list-style-type: none"> <li>• Praktische redenen</li> <li>• Het screenen met vaste koppels brengt risico’s met zich mee: er ontstaan structuren / patronen waardoor afwijkingen structureel kunnen worden gemist.</li> </ul>
I	Nee	n.v.t.	n.v.t.	Praktisch niet uitvoerbaar.
J	Nee	n.v.t.	n.v.t.	<ul style="list-style-type: none"> <li>• De diversiteit van de individueel bestaande screeningstrategieën benutten voor een optimaal effect op de VW en DT cijfers van de totale maatschap.</li> <li>• Individuele wensen aangaande screeningsdagen en tijdstippen</li> </ul>

	Vraag 1	Vraag 2	Vraag 3	Vraag 4
BE	<b>Wordt er binnen de BE (deels) gescreend in vaste koppels?</b>	<b>Zo ja, hoeveel koppels op het totaal?</b>	<b>Zo ja, worden (individuele) resultaten van verwijscijfers en detectiecijfers gebruikt om deze koppels samen te stellen?</b>	<b>Zo nee, graag beschrijven welke redenen er zijn om niet in vaste koppels te screenen:</b>
K	Nee bij voorkeur niet. (In het verleden wel gedaan. ) Maar aangezien screeners op vast dagen komen ontstaan er meerdere koppels.	Een koppel kan ontstaan met een screener van de dag ervoor, zelfde dag of dag erna.	Nee, individuele verwijs- en detectiecijfers zijn wel voor de hele groep bekend. Er wordt rekening gehouden met elkaars kracht en mindere kanten	Nadeel: je groeit naar elkaar toe. Juist diversiteit in koppels blijft je lerend vermogen alert houden. Iedereen in de groep vindt variatie en terugkoppeling belangrijk.
M	Nee	n.v.t.	n.v.t.	Roostertechnisch niet haalbaar i.v.m. ons multi-locatie systeem .
P	Nee	n.v.t.	n.v.t.	Geen voordeel vast koppels. Dit is gemakkelijker plannen.
Q	Nee	n.v.t.	n.v.t.	<ul style="list-style-type: none"> <li>• Plantechisch niet mogelijk.</li> <li>• Mix van radiologen lijkt ons niet verkeerd; leereffect van elkaars verwijzingen</li> </ul>

## Resultaten: Omvang sessie

Vraag 6					
BE	Hoeveel dagproducties worden er (gemiddeld genomen) per sessie gelezen door een screeningsradioloog?	Verwijs /1000 vervolgscr.	Detectie /1000 vervolgscr.	PVW [%] vervolgscr.	Sens [%] totaal
B	Gemiddeld 3	28,1	6,6	23,5	76,4
C	1	22,7	6,8	29,7	74,3
D	3 dagproducties	19,8	5,7	28,7	-
E	Tussen de 4 en de 8 dagproducties	17,4	5,5	31,6	69,0
F	4 SE's per dag. De screeningsradiologen werken een dagproductie in één sessie af.	13,0	5,8	44,4	81,4
G	5	21,4	6,7	31,3	74,1
H	Alleen 's avond screenen, dan 3. En voor degene die overdag screenen 6.	19,7	6,3	32,0	82,6
I	4 dagproducties	18,4	6,0	32,8	73,2
J	300 cliënten (4 -5 dagproducties)	18,5	6,0	32,5	76,7
K	gemiddeld 300 - 350 cliënten (ca. 5 dagproducties)	20,3	7,0	34,4	76,8
M	2	18,8	6,5	34,6	80,5
P	Meestal 5	14,8	6,2	42,0	78,2
Q	gemiddeld 250 cliënten (3 -4 dagproducties)	16,5	5,3	32,5	72,1

## Resultaten: Double reading

Vraag 5		Vraag 7				
BE	Wordt er bij het inplannen rekening gehouden met de verdeling 1 <sup>e</sup> en 2 <sup>e</sup> lezingen van een screeningsradioloog?	Wordt er volledig blind gelezen, of wordt er op een of andere manier informatie aan elkaar doorgegeven? Zo ja, dan graag beschrijven op welke wijze:	Verwijs/ 1000 vervolgscr.	Detectie /1000 vervolgscr.	PVW [%] vervolgscr.	Sens [%] totaal
B	Nee. Bij dubbelblind lezen maakt 1 <sup>e</sup> of 2 <sup>e</sup> lezing geen verschil. Niet ervaren screeningsradioloog arbitreert niet.	Volledig dubbelblind	28,1	6,6	23,5	76,4
C	Wij proberen dat wel, maar de combinatie met werk in het ziekenhuis maakt dat vaak moeilijk.	Er wordt volledig blind gelezen, maar we maken wel annotaties over de verwezen cliënten om de volledigheid van het verwijsproces door de screeningorganisatie te monitoren.	22,7	6,8	29,7	74,3
D	Nee, omdat wij de arbitrage laten doen door de 3 <sup>e</sup> lezer.	Volledig blind	19,8	5,7	28,7	-
E	Iedereen leest datum gerelateerd eerst de 1 <sup>e</sup> lezingen en daarna de 2 <sup>e</sup> lezingen met een homogene verdeling 1 <sup>e</sup> /2 <sup>e</sup> lezing. Tevens elke dag afhandeling van de arbitrages.	Volledig blind	17,4	5,5	31,6	69,0
F	Nee. Het aantal screeningsonderzoeken per radioloog is dermate hoog ( $\pm 15.000$ /jr) dat een ieder ruim voldoende aan bod komt als 1 <sup>e</sup> of 2 <sup>e</sup> screener, ook al zijn er wat dit betreft verschillen tussen radiologen.	Er wordt vooraf geen inhoudelijke informatie uitgewisseld m.b.t. individuele screeningsonderzoeken. Wel wordt meestal door de 1 <sup>e</sup> lezer aan de 2 <sup>e</sup> lezer mondeling doorgegeven hoeveel onderzoeken beoordeeld moeten worden. Soms volgt dan nog een algemene opmerking over de moeilijkheidsgraad van de betreffende dagproductie (b.v. 'veel Bi-Rads 0 of juist 'veel Bi-Rads 4/5' of 'veel (onzinnige;) verwijzingen door de laboranten' o.i.d.)	13,0	5,8	44,4	81,4



Vraag 5		Vraag 7					
BE	Wordt er bij het inplannen rekening gehouden met de verdeling 1 <sup>e</sup> en 2 <sup>e</sup> lezingen van een screeningsradioloog?	Wordt er volledig blind gelezen, of wordt er op een of andere manier informatie aan elkaar doorgegeven? Zo ja, dan graag beschrijven op welke wijze:	Verwijs/ 1000 vervolgscr.	Detectie /1000 vervolgscr.	PVW [%] vervolgscr.	Sens [%] totaal	
G	Ja, de bedoeling is dat aan het eind van elk kalenderjaar elke screeningsradioloog ongeveer hetzelfde aantal beoordelingen heeft gedaan en dat er een ongeveer een gelijke verdeling is in 1 <sup>e</sup> en 2 <sup>e</sup> lezingen.	Dubbelblind, er wordt een lijst bijgehouden van de vewezen cliënten door de 1 <sup>e</sup> lezer, deze is niet inzichtelijk voor de 2 <sup>e</sup> lezer tijdens het beoordelen, alleen achteraf ter controle.	21,4	6,7	31,3	74,1	
H	Nee, niet speciaal, maar we proberen binnen een jaar ongeveer 50% 1 <sup>e</sup> en 2 <sup>e</sup> lezingen te doen.	Volledig blind	19,7	6,3	32,0	82,6	
I	Nee; moeilijk uitvoerbaar, maar ook niet noodzakelijk; blijkt in de praktijk aardig fifty fifty verdeeld.	Volledig blind	18,4	6,0	32,8	73,2	
J	Nee	Lezer 1 noteert op een papieren werklijst van de betreffende SE zowel de redenen van zijn/haar verwijzing met BIRADS codering, alsook bij andere gesignaleerde afwijkingen de redenen om niet te verwijzen. Lezer 2 heeft de keuze: eerst de lezing doen en leest volledig blind of neemt vooraf kennis van de mededelingen van lezer 1 op de papieren werklijst en leest daarna de productie.	18,5	6,0	32,5	76,7	
K	In onze groep wordt het aantal lezingen per week evenredig verdeeld over de screeners in de groep. De screener zorgt zelf voor een gelijk aantal 1 <sup>e</sup> en 2 <sup>e</sup> lezingen (per week). Alle screeners in onze groep zijn gelijk (geen junior/senior). Elke maand krijgen de screeners te zien wat het individuele % 1 <sup>e</sup> en 2 <sup>e</sup> lezingen is.	Volledig blind	20,3	7,0	34,4	76,8	
M	Ja, maar niet leidend. Zie ook vraag 8.	Volledig blind	18,8	6,5	34,6	80,5	
P	Nee maar de mix blijkt vanzelf goed	Volledig blind	14,8	6,2	42,0	78,2	
Q	Nee	Volledig blind	16,5	5,3	32,5	72,1	

## Resultaten: Omgaan met discrepanties

Vraag 8						
	Hoe wordt er om gegaan met discrepanties? Graag de procedure beschrijven:	Methode*	Verwijs/ 1000 vervolgscr.	Detectie /1000 vervolgscr.	PVW [%] vervolgscr.	Sens [%] totaal
B	Alle discrepante readings worden beoordeeld door een derde reader (welke niet de eerste of de tweede reader betreft).	5	28,1	6,6	23,5	76,4
C	Bij discrepanties volgt in eerste instantie overleg tussen radioloog 1 en radioloog 2. In geval van BI-RADS 0 discrepante verwijzingen mag dit leiden tot zowel een uitkomst BI-RADS 0 als BI-RADS 1. In het geval van een discrepante verwijzing BI-RADS 4 of 5 mag de niet-verwijzende screeningsradioloog zijn initiële codering upgraden naar BI-RADS 4 of 5. Downcoding is echter niet wenselijk. Blijft de niet-verwijzende screeningsradioloog bij zijn oorspronkelijke oordeel, dan is daarom arbitrage door een derde voorgeschreven. Omdat de volgende werkdag ten minste één andere screeningsradioloog is ingedeeld, zal deze de arbitrage direct de volgende werkdag kunnen verrichten.	3	22,7	6,8	29,7	74,3
D	Arbitrage door 3 <sup>e</sup> lezer.	5	19,8	5,7	28,7	-
E	Indien 1 <sup>e</sup> en 2 <sup>e</sup> lezer een discrepantie hebben, komt het onderzoek direct op de arbitragelijst. Vervolgens komt er een 3 <sup>e</sup> radioloog aan te pas, die de definitieve beslissing neemt.	5	17,4	5,5	31,6	69,0
F	2e lezer als 3e lezer. Met daarmee 2 mogelijkheden: óf de 2e lezer verwijst alsnog door, óf de cliënt wordt op een arbitragelijst gezet	1C	13,0	5,8	44,4	81,4
G	Systeem van arbitrages: 1 <sup>e</sup> en 2 <sup>e</sup> reader geen concensus, 3 <sup>e</sup> reader (niet zijn de de 1 <sup>e</sup> of 2 <sup>e</sup> reader) geeft het finale oordeel.	5	21,4	6,7	31,3	74,1
H	Als de 2 <sup>e</sup> lezer zich kan aanpassen aan de 1 <sup>e</sup> wordt de casus direct afgewerkt. In alle andere gevallen wordt de casus op de arbitrage lijst gezet en kijkt een 3 <sup>e</sup> radioloog.	1C	19,7	6,3	32,0	82,6
I	In geval van een discrepantie wordt er gearbitreerd door een 3e screeningsradioloog	5	18,4	6,0	32,8	73,2

Vraag 8						
Hoe wordt er om gegaan met discrepanties? Graag de procedure beschrijven:		Methode*	Verwijs/ 1000 vervolgscr.	Detectie /1000 vervolgscr.	PVW [%] vervolgscr.	Sens [%] totaal
J	Wanneer reader 2 een niet door reader 1 gewenste verwijzing toch wil doorverwijzen, mag deze zonder overleg dit als derde reader doen. Indien reader 2 een verwijzing van reader 1 niet wil volgen, dan komt deze cliënt op de arbitragelijst en wordt door 1 van de readers uit het team van de andere dag als derde beoordeeld	1C	18,5	6,0	32,5	76,7
K	Alle discrepanties worden doorgezet naar arbitrage (derde radioloog). Getracht wordt om van de arbitrages slechts 30% te verwijzen (uit eerder onderzoek hebben we gezien dat slechts 8% van alle discrepanties TP is!). Uitzondering: Alleen in geval van OEPS mag een discrepantie direct door de 2e reader verwezen worden. OEPS geval is een situatie waarin een 2e reader iets significant niet gezien heeft en wel wil verwijzen. Dit om geen vertraging op te lopen.	1C	20,3	7,0	34,4	76,8
M	Via consensus. Beide lezers zijn fysiek aanwezig op moment van het verwerken van de discrepanties.	3	18,8	6,5	34,6	80,5
P	Altijd derde radioloog als eerste actie van de volgende dag.	5	14,8	6,2	42,0	78,2
Q	bij discrepanties beoordeeld een derde radioloog samen met de tweede lezer. Deze komen tot consensus en bepalen of cliënt wel of niet wordt doorverwezen	2	16,5	5,3	32,5	72,1

- \* In de LRCB notitie n.a.v. het proefschrift van Dr. E.G. Klompenhouwer worden vijf verschillende **methoden** beschreven voor het oplossen van discrepante beoordelingen:
- De 2<sup>e</sup> lezer beoordeelt opnieuw met de informatie van de beoordeling van de 1<sup>e</sup> lezer (2<sup>e</sup> lezer = 3<sup>e</sup> lezer), eventueel gevolgd door arbitrage door onafhankelijke lezer. Hierbij wordt een onderverdeling gemaakt waarbij:
    - de 2<sup>e</sup> lezer mag BI-RADS 0, 4 en 5 downgraden.
    - de 2<sup>e</sup> lezer mag alleen een BI-RADS 0 downgraden.
    - de 2<sup>e</sup> lezer mag in geen enkel geval downgraden.
  - De 2<sup>e</sup> lezer beoordeelt opnieuw met de informatie van de beoordeling van de 1<sup>e</sup> lezer, samen met een onafhankelijke 3<sup>e</sup> lezer (niet zijnde de 1<sup>e</sup> lezer). De 2<sup>e</sup> lezer en 3<sup>e</sup> lezer komen tot consensus.
  - De 1<sup>e</sup> en 2<sup>e</sup> lezer komen tot consensus. Indien geen overeenstemming bereikt kan worden, dan volgt arbitrage door een 3<sup>e</sup> onafhankelijke lezer.
  - Er wordt altijd verwezen.
  - Er wordt altijd een 3<sup>e</sup> onafhankelijke lezer gevraagd (arbitrage).

## Resultaten: Hanging protocol

BE	Vraag 9	Vraag 10
	<b>Welk hanging protol wordt er gebruikt?</b> a) Het door de fabrikant ingestelde hanging protocol b) Het door de screeningsorganisatie ingestelde hanging protocol c) Het door de groep/maatschap ingestelde hanging protocol d) Alle screeningsradiologen van de groep hebben een persoonlijk ingesteld hanging protocol e) Onbekend	<b>Wordt de toggle/blinken functie door de screeningsradiologen gebruikt?</b> a) Door iedereen b) Door het grootste gedeelte van de groep c) Door een enkeling van de groep d) Door niemand van de groep e) Onbekend
B	a)	a)
C	a)	c)
D	b)	b)
E	c)	a)
F	c) en d)	a)
G	c) en d)	a)
H	e)	a)
I	d)	e)
J	d)	e)
K	d)	a)
M	De meesten b), kleiner aantal d)	c)
P	c) en d)	a)
Q	b)	c)

## Discussie en conclusie

### *Screenen in koppels*

In haar proefschrift doet Dr. Klompenhouwer de suggestie om specifieke radiologenkoppels samen te stellen: een screeningsradioloog met een zeer goede performance koppelen aan een radioloog met mindere performance of een radioloog die niet gespecialiseerd is in breast imaging koppelen met een radioloog die dat wel is. Ze schrijft ook dat dit in de praktijk lastig kan zijn (pagina's 47 en 113 proefschrift).

Het LRCB onderschrijft het idee om speciale radiologenkoppels samen te stellen. Momenteel worden beginnende screeningsradiologen al geadviseerd om na het afronden van het 8-daagse inwerkprogramma te lezen als koppel samen met een ervaren collega screeningsradioloog. Voor zover het LRCB daar inzage in heeft, wordt dit in de praktijk ook zo gedaan.

Uit de inventarisatie blijkt dat er momenteel in Nederland niet in koppels gescreend wordt. Als reden om dit niet te doen wordt het vaakst genoemd dat dit praktisch niet uitvoerbaar is. Vier BE's zien de diversiteit door niet in vaste koppels te screenen als een voordeel, als een leereffect. Door twee BE's wordt de angst uitgesproken dat er patronen kunnen ontstaan in vaste koppels, waardoor afwijkingen gemist zouden kunnen worden. De conclusie is dat er bij de BE's momenteel geen draagvlak is om in vaste koppels te gaan screenen. Het LRCB ziet op dit moment geen aanleiding om dit punt verder te belichten.

### *Omvang sessie*

Uit de inventarisatie blijkt dat er erg veel variatie is qua omvang van een sessie die per keer door een screeningsradioloog wordt beoordeeld, zowel tussen de verschillende BE's als binnen een BE.

Het LRCB ziet wel een meerwaarde om hier verder onderzoek naar te doen. Er zijn drie vragen waar het LRCB graag een antwoord op zou willen krijgen:

- Wat is de optimale omvang van een sessie?
- Heeft de volgorde van lezen van de sessie invloed op de performance?
- Zouden radiologen moeten "opwarmen" voordat ze aan een sessie beginnen?

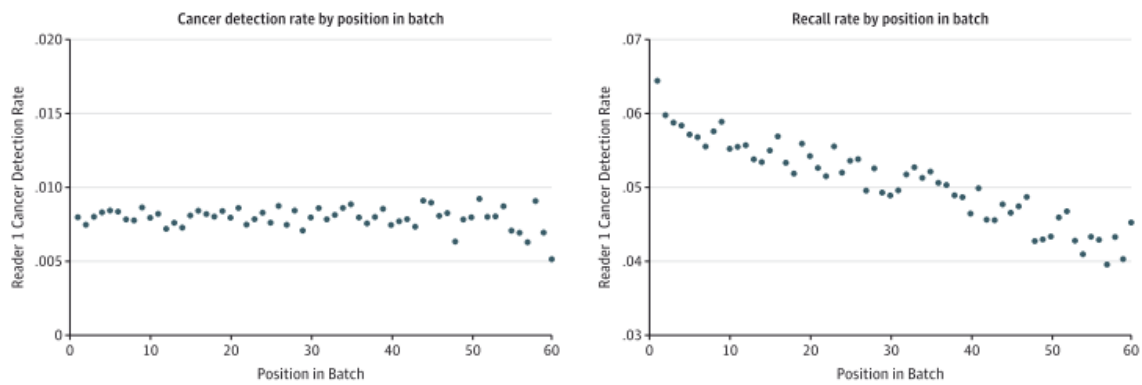
In 2016 is er een artikel verschenen van S. Taylor-Phillips et. al.<sup>1</sup>, over het effect van de volgorde van lezen. Dit onderzoek is uitgevoerd binnen de Engelse borstkankerscreening. De conclusie van dit onderzoek was dat er geen significant verschil gemeten werd in verwijs- en detectiecijfer als de 1<sup>e</sup> en 2<sup>e</sup> lezer dezelfde volgorde lazen of de 1<sup>e</sup> en 2<sup>e</sup> lezer een tegengestelde volgorde lazen.

---

<sup>1</sup> Taylor-Phillips S, MG W, Jenkinson D, Al E (2016) Effect of using the same vs different order for second readings of screening mammograms on rates of breast cancer detection: A randomized clinical trial. JAMA 315:1956–1965. doi: 10.1001/jama.2016.5257

Een ander opvallend resultaat van deze studie was dat gedurende de leestijd, gemiddeld over heel veel sessies, de gemiddelde detectie constant blijft, maar dat het gemiddelde verwijscijfer afneemt gedurende de leestijd. Zie onderstaand figuur uit het artikel:

Figure 3. Average Patterns of Cancer Detection Rate and Recall Rate for a Single Reader Over the Course of Examining a Batch of Mammograms



Each data point represents the mean recall or cancer detection rate over all cases examined by reader 1 at that position in the batch. A total of 1 173 930

cases were included, examined as reader 1 by 348 readers. The median number of screenings per batch position is 21 931 (interquartile range, 10 133-28 126).

Hierdoor is bij ons de vraag opgekomen of radiologen, net als sporters, zouden moeten “opwarmen” voordat ze gaan screenen. Het onderzoek van Taylor zou met gebruik van de Nederlandse data relatief makkelijk herhaald kunnen worden.

Eerder heeft het LRCB een literatuuronderzoek uitgevoerd, met de conclusie dat er uit de literatuur niet duidelijk kan worden vastgesteld wat het maximaal aantal uren per dag is dat een screeningsradioloog verantwoord kan beoordelen. Het nemen van regelmatige pauzes is belangrijk om oogspanning en vermoeidheid te verminderen en screeningsradiologen dienen zich hier dan ook bewust van te zijn.

### **Double reading**

Uit de inventarisatie blijkt dat het merendeel van de BE's (8) bij de planning geen rekening houdt met de verdeling 1<sup>e</sup> en 2<sup>e</sup> lezingen van een screeningsradioloog. Hieruit trekt het LRCB de conclusie dat de screeningsradiologen mogelijk niet voldoende het belang inzien van een evenredige verdeling. Het LRCB vindt het belangrijk dat een screeningsradioloog even vaak 1<sup>e</sup> lezer is als 2<sup>e</sup> lezer, omdat een radioloog als 2<sup>e</sup> lezer meer verantwoordelijkheid draagt dan als 1<sup>e</sup> lezer:

- de 2<sup>e</sup> lezer doet de verslaglegging;
- de 2<sup>e</sup> lezer kent de uiteindelijke BI-RADS toe;
- de 2<sup>e</sup> lezer heeft geen mogelijkheid meer zijn mening te herzien.

Indien de discrepanties worden opgelost volgens methode 1C (zie omgaan met discrepanties), is er voor de 2<sup>e</sup> lezer tevens een leereffect.

Verder blijkt uit deze inventarisatie dat bij het merendeel van de BE's (11) volledig blind wordt gelezen. Uit de studies van Dr. Klompenhouwer bleek dat volledig blind lezen een betere programmasensitiviteit geeft dan niet blind lezen (pagina 62), maar wel met een veel hoger verwijscijfer. Uit de data van de visitaties volgt dat er veel verschil is tussen de BE's, maar uit deze data kan niet de conclusie getrokken worden dat de BE's die niet volledig blind lezen (G en J) een lagere programmasensitiviteit hebben. Even goed gaat de voorkeur van het LRCB uit naar volledig blind lezen. Het is namelijk niet ondenkbaar dat het een bepaald "gemak" zou kunnen veroorzaken als je niet volledig blind leest.

### ***Omgaan met discrepanties***

Uit de inventarisatie volgt dat de helft van de BE's (6) in 100% van de gevallen een discrepantie oplossen doormiddel van arbitrage door een 3<sup>e</sup> onafhankelijke lezer (methode 5). De daarop volgende meest gebruikte methode (4 BE's) is dat de 2<sup>e</sup> lezer opnieuw beoordeelt met de informatie van de beoordeling van de 1<sup>e</sup> lezer (2<sup>e</sup> lezer = 3<sup>e</sup> lezer), waarbij in geen enkel geval de BI-RADS mag worden gedowngrade. Met andere woorden, indien de 2<sup>e</sup> lezer bij zijn mening blijft niet te willen verwijzen, dan volgt een arbitrage door een onafhankelijke lezer (methode 1C). De overige 3 BE's hanteren een methode op basis van consensus: consensus tussen 1<sup>e</sup> en 2<sup>e</sup> lezer (methode 3) of consensus tussen 2<sup>e</sup> lezer en onafhankelijke 3<sup>e</sup> lezer (methode 2).

In haar studie heeft Dr. Klompenhouwer een vergelijking tussen blinded double reading met en zonder arbitrage (tabel 4.3.1, pagina 84) uitgevoerd. Haar conclusie is dat arbitrage als positief effect heeft dat het verwijscijfer daalt en de PVW toeneemt, maar als negatief effect heeft dat de programmasensitiviteit afneemt. Daarom schrijft ze in haar proefschrift dat ze arbitrage geen goed middel vindt voor het Nederlandse bevolkingsonderzoek (zie pagina 88 van het proefschrift). In een vervolg studie (zie hoofdstuk 4.4 van het proefschrift) heeft Dr. Klompenhouwer onderzocht of arbitrage wel een goed middel zou zijn, indien deze alleen wordt toegepast op BI-RADS 0 discrepanties in plaats van op alle discrepanties. Ze komt tot een positieve conclusie: ook hier heeft de arbitrage het positieve effect dat het verwijscijfer daalt en de PVW toeneemt. Maar nu zonder effect op de programmasensitiviteit (pagina 104). In haar proefschrift adviseert zij dan ook arbitrage van discrepante BI-RADS 0 verwijzingen (pagina 106).

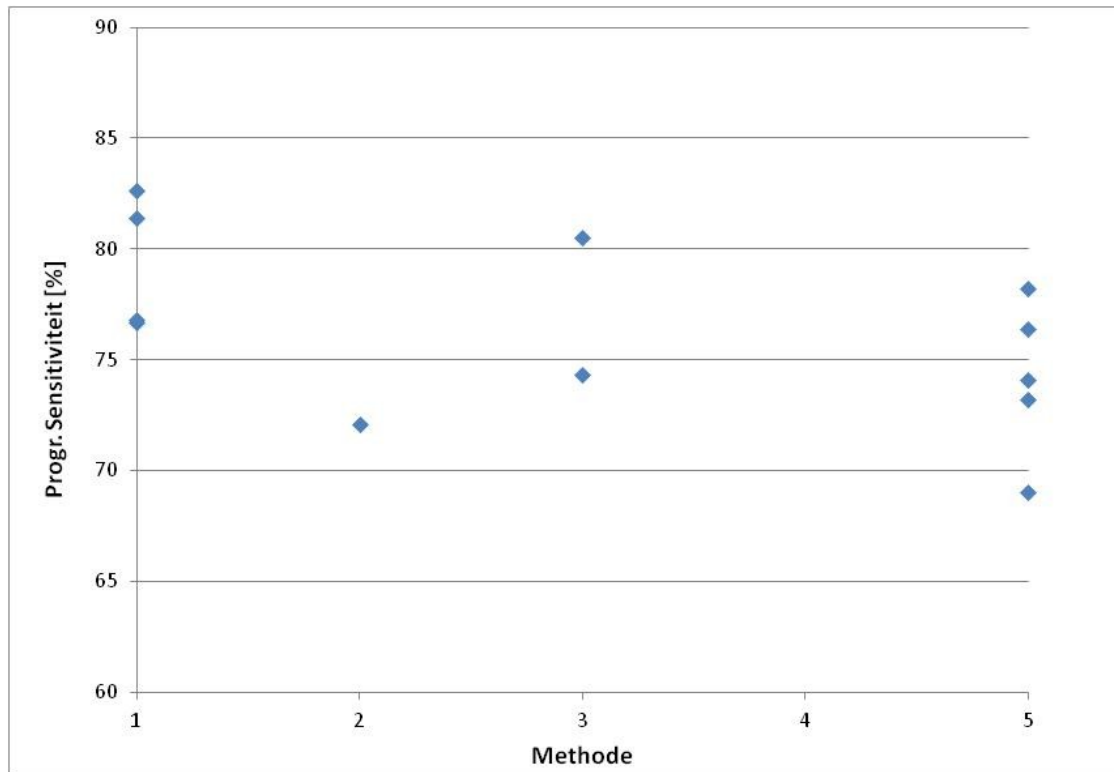
Dr. Klompenhouwer heeft in haar studies niet gekeken naar discrepanties oplossen doormiddel van consensus. Colette M. Staw en haar collega's hebben wel een studie uitgevoerd naar het effect van consensus (uitgevoerd binnen de Ierse nationale borstkankerscreening)<sup>2</sup>. Hier is de conclusie eigenlijk gelijk: het voordeel van een

---

<sup>2</sup> Shaw CM, Flanagan FL, Fenlon HM, McNicholas MM. Consensus review of discordant findings maximizes cancer detection rate in double-reader screening mammography: Irish National Breast Screening Program experience. *Radiology*. 2009;250(2):354–362.

consensus meeting t.o.v. alle discrepanties doorverwijzen is dat het verwijscijfer daalt, maar dat de sensitiviteit ook iets daalt.

Er is geen studie bekend die de methoden van consensus en arbitrage met elkaar vergelijkt. Maar als de sensitiviteit van de BE's (data uit de visitaties) wordt uitgezet tegen de gebruikte methode voor het oplossen van discrepanties (zie grafiek hieronder), lijkt het erop dat de methode van arbitrage een iets lagere sensitiviteit geeft dan de andere methoden. Het LRCB onderschrijft de bevinding van Dr. Klompenhouwer dat arbitrage niet voor alle discrepanties ingezet moet worden.



Figuur 1: Programmasensitiviteit van de BE's (data uit de visitaties) uitgezet tegen de gebruikte methode voor het oplossen van discrepanties

Gevoelsmatig wordt ook een leermoment gemist als direct overgegaan wordt op arbitrage. Bij een consensus tussen twee lezers vindt er een discussie plaats over een moeilijke casus, wat gezien mag worden als leermoment. Dit wordt ook zo beschreven in het artikel van Shaw et al.: "In comparison, consensus review at our center enabled consultants to discuss and tease out the nature of the abnormality rather than proceed with further independent blinded reviews.....This technique provided readers with a forum for discussion of difficult cases and served as an educational tool for both consultant radiologists and residents."

Ook bij methode 1C is er sprake van een leermoment. Je bekijkt je eigen mogelijke missers opnieuw en maakt een heroverweging.

Als op dit moment, met de huidige informatie, uit het oogpunt van een uniforme werkwijze een keuze gemaakt zou moeten worden uit de verschillende methoden van oplossen van discrepanties, geeft het LRCB de voorkeur aan methode 1C. Dit



geeft een leereffect voor de 2<sup>e</sup> lezer en kan ook worden uitgevoerd door BE's waar het praktisch niet mogelijk is om twee lezers fysiek met elkaar in discussie te laten gaan.

Het LRCB is van mening dat er op dit moment nog wel een aantal vragen liggen die verdere aandacht verdienen:

- Er zou een discussie gevoerd moeten worden over waar de nadruk op zou moeten liggen binnen de Nederlandse borstkankerscreening: op een hogere programma sensitiviteit (met als nadeel meer fout-positieven) of juist een verlaging van het aantal fout-positieven met een mogelijk lagere programma sensitiviteit.
- Er zou onderzocht kunnen worden welk percentage van de verwijsadviezen tot stand is gekomen na een discrepantie. Tevens zou onderzocht kunnen worden wat voor soort laesies er zitten in deze verwijsadviezen in vergelijking met de laesies in de verwijsadviezen waarbij 1<sup>e</sup> en 2<sup>e</sup> lezer direct zouden verwijzen. Verwacht wordt dat de verwijsadviezen na een discrepantie meer DCIS en kleinere invasieve tumoren zullen bevatten.
- Er zou bij de intervalcarcinomen onderzocht kunnen worden, bij welk percentage er een discrepantie was en er uiteindelijk geen verwijsadvies volgde.
- Dr. Klompenhouwer adviseert n.a.v. haar studie blinded double reading met arbitrage voor alleen de BI-RADS 0. De studie van Dr. Klompenhouwer is uitgevoerd in één regio, in een onderzoeksgroep met enthousiaste radiologen. Daarmee is niet duidelijk of de resultaten reproduceerbaar zijn binnen andere regio's. Het onderzoek zou herhaalt kunnen worden in een andere regio.
- De kosteneffectiviteit van de verschillende methoden zou onderzocht kunnen worden.

### ***Hanging protocol***

Uit de inventarisatie blijkt dat er landelijk veel verschillen zijn in het gebruikte hanging protocol, zelfs binnen één BE. Verder blijkt dat er door het merendeel van de screeningsradiologen gebruik gemaakt wordt van de toggle functie.

Het LRCB heeft geen voorkeur voor het ene hanging protocol boven het andere. Wel is het LRCB van mening dat van twee methoden intensief gebruik gemaakt zou moeten worden: van de toggle functie (vanwege perceptie) en van het overzicht oud/nieuw (vanwege het afstand nemen). Tijdens de herbeoordelingen bij de visitaties is met enige regelmaat gebleken dat deze twee methoden verduidelijking gaf.

## **Samenvatting**

Opgrond van de huidige informatie heeft het LRCB een voorkeur voor de volgende leesstrategie:

- Volledig blind lezen;
- Methode 1C voor het oplossen van discrepanties;
- Evenredige verdeling van 1<sup>e</sup> en 2<sup>e</sup> lezingen door een screeningsradioloog;
- Intensief gebruik maken van de toggle functie en het overzicht oud/nieuw.

Punten die verder belicht zouden moeten worden:

- Invloed van de omvang sessie op de performance;
- Nadruk op een hogere programma sensitiviteit of juist een verlaging van het aantal fout-positieven;
- Aandeel verwijsadviezen met een discrepante beoordeling;
- Aandeel intervalcarcinomen met een discrepante beoordeling.