Review

Considerations for Evaluating the Introduction of New Cancer Screening Technology: Use of Interval Cancers to Assess Potential Benefits and Harms

Rachel Farber [1,*], Nehmat Houssami [1,2], Isabelle Barnes [3,4,5], Kevin McGeechan [1], Alexandra Barratt [1] and Katy J. L. Bell [1]

1   Wiser Healthcare, Sydney School of Public Health, Faculty of Medicine and Health, The University of Sydney, Sydney 2006, Australia
2   The Daffodil Centre, The University of Sydney, A Joint Venture with Cancer Council NSW, Sydney 2006, Australia
3   Sydney School of Public Health, Faculty of Medicine and Health, The University of Sydney, Sydney 2006, Australia
4   Centre for Women's Health Research, College of Health, Medicine and Wellbeing, The University of Newcastle, Callaghan 2308, Australia
5   Australian Longitudinal Study on Women's Health, The University of Newcastle, Callaghan 2308, Australia
*   Correspondence: rachel.farber@sydney.edu.au

Abstract: This framework focuses on the importance of the consideration of the downstream intermediate and long-term health outcomes when a change to a screening program is introduced. The authors present a methodology for utilising the relationship between screen-detected and interval cancer rates to infer the benefits and harms associated with a change to the program. A review of the previous use of these measures in the literature is presented. The framework presents other aspects to consider when utilizing this methodology, and builds upon an existing framework that helps researchers, clinicians, and policy makers to consider the impacts of changes to screening programs on health outcomes. It is hoped that this research will inform future evaluative studies to assess the benefits and harms of changes to screening programs.

Keywords: cancer; cohort; ductal carcinoma in situ; mammography; recall; breast mammography; digital; health outcomes; film screen; interval cancer

1. Introduction

Cancer screening programs aim to reduce morbidity and mortality through the early detection and treatment of cancers that would have otherwise have gone on to present at a later stage [1]. When a change is introduced to a screening program, such as a change in technology, it is important to evaluate whether this change results in an overall net benefit. The potential benefits from the change should not look only at the initial impact on improved sensitivity, but also evaluate whether this improved sensitivity addresses clinically important cancers where early detection is, providing outcomes which are likely to translate into longer-term benefits [2]. While it is important to estimate the increases in screen-detected cancers, we need to know the extent to which the additional cancers detected are those that would have otherwise progressed to present later, representing potential beneficial early detection—versus indolent lesions that would not have progressed at all, or would only have grown very slowly—representing potential harm from over-detection.

2. Role of Interval Cancers

The prompt assessment of the effects of new technology or screening protocols requires methods using short-term outcomes [3]. Once a screening program is underway, interval cancers can be used as a proxy measure for cancers not currently being detected by the

Int. J. Environ. Res. Public Health **2022**, *19*, 14647

2 of 17

existing screening program [4,5]. Interval cancers are cancers diagnosed after a negative screening result and before a subsequent scheduled screening. Interval cancers may not be detected during screening either because they: did not yet exist, were not detectable with the modality used or were missed by the person reading the screens. Most interval cancers are symptomatic and are therefore by definition clinically important. They have been found to have similar prognostic features to other cancers clinically diagnosed outside of screening [6]. Interval cancers tend to be more biologically aggressive tumours with a faster growth rate and more unfavourable tumour characteristics than screen-detected cancers [7–9]. Therefore, interval cancers represent the cancers where early detection would be beneficial.

For a change to a screening program to increase its effectiveness, it should result in an increase in the diagnosis of clinically important early stage cancers, a decrease in interval cancers diagnosed between scheduled screenings, and ultimately a decline in the incidence of cancers diagnosed at late stage. However, if such changes are not seen, any additional cancers diagnosed via a change to the program may represent overdiagnosed cancers that would not have otherwise presented, or more slowly growing cancers that would have been detected in the old screening program without detriment [3,10]. The emphasis in screening should not be on finding more cancers, but on finding more cancers that are clinically important [11]. The tension between maximising benefit through the detection of clinically important cancers and minimising harm through the detection of clinically unimportant lesions, means that screening programs ought to undergo regular re-evaluation, particularly when changes are introduced [12,13]. It is important to ensure that any changes in the implementation of screening optimize the balance between benefit and harm, including minimising the risk of overdiagnosis [14].

3. Relationship between Screen-Detected and Interval Cancers

A reduction in interval cancers is sometimes used as a surrogate measure of a change in screening effectiveness [5]. Previously, screen-detected cancer rates and interval cancer rates have been used to measure the benefit from a new intervention. However, it is also important to consider the reverse implication of an increase in detection rate without a meaningful decrease in interval cancer rates. Our framework takes the relationship between interval cancers and screen-detected cancers a step further, to consider how this relationship might be used as both an indicator of benefit, and of harm, from changes made to the program.

Figure 1 shows the possible scenarios concomitant upon a change in technology. If a change to a new technology increases sensitivity for clinically important cancers we expect to see:

(1) an increase in screen-detected cancers with tumour markers indicating aggressive disease; and

(2) a decrease in the interval cancer rate between screenings, indicating that at least some of the additional detection is occurring in clinically important and rapidly progressing cancers.

However, if the change only increases the detection of clinically unimportant cancers, we expect to see:

(3) an increase in the detection of screen-detected cancers with tumour markers indicating low risk disease; and

(4) no or minimal changes in the interval cancer rate between screenings, indicating that the additional detection is of overdiagnosed or slow progressing cancers.
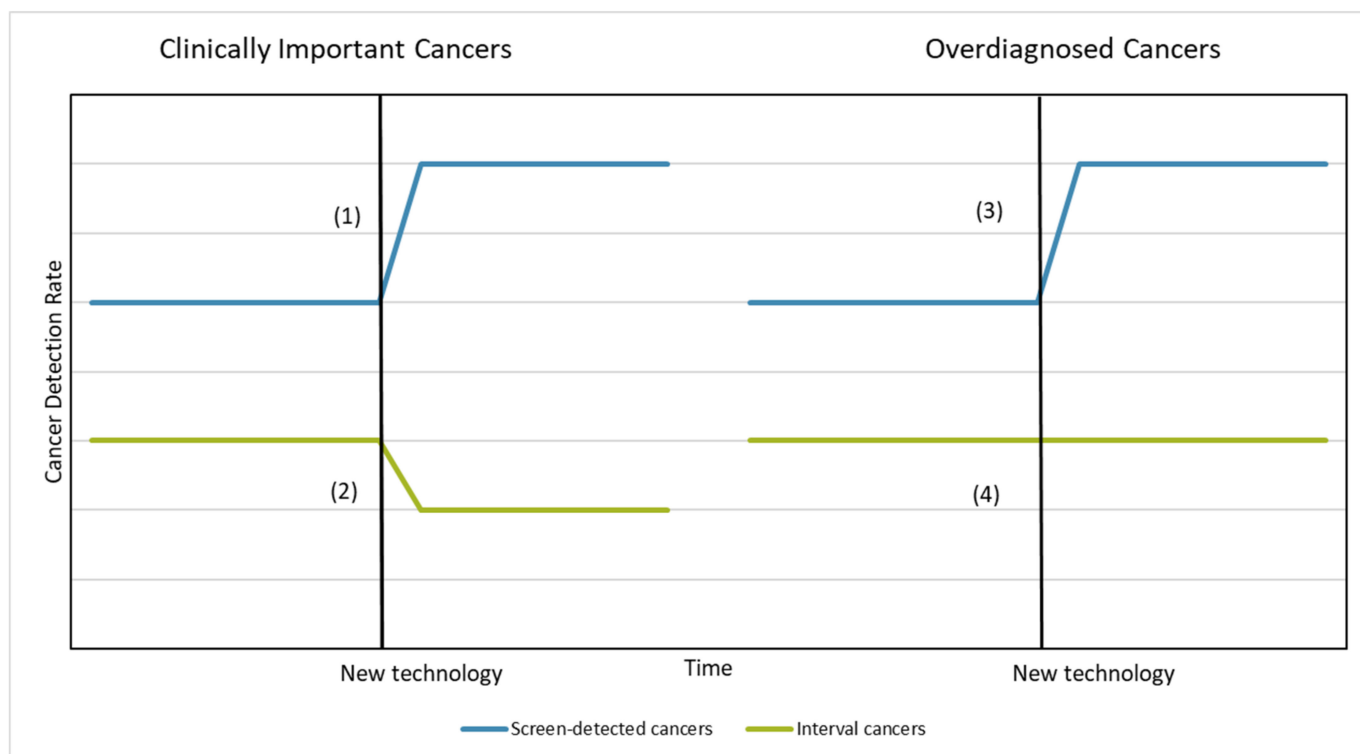
Figure 1. Potential scenarios following the introduction of a new technology to a screening program.

The relationship between screen-detected and interval cancers can be further identified by plotting the risk difference in screen-detected and interval cancer rates before and after the change. Rates are typically presented per 1000 screens.

In Figure 2 each quadrant represents a different signature:

(1) The top left quadrant shows a decrease in screen detection and an increase in interval rates. This observation indicates that there is decreased rate of detection, including clinically important cancers, and therefore suggests an increase in underdiagnosis

(2) The bottom left quadrant shows a decrease in screen detection and a decrease in interval rates. This result indicates that there is decreased detection overall, but an increased detection of clinically important cancers, and therefore suggests a decrease in overdiagnosis. Theoretically this could occur with risk-stratified screening [15].

(3) The top right quadrant shows an increase in screen detection and no change, or possibly even an increase, in interval rates. Assuming no change in the background rates of cancer, this observation would indicate that the increased detection represents an increase in overdiagnosis.

(4) The bottom right quadrant shows an increase in screen detection and a decreased interval rate. This finding indicates that the increased detection does include clinically important and progressing cancers, and therefore suggests a decrease in underdiagnosis. The second and fourth quadrants show the improved effectiveness sought with a change to the screening program.

It is important to look not only at the relationship between screen-detected and interval cancer rates, but also at the tumour characteristics of the cancers detected. Differences between technologies in both screen-detected cancers and interval cancers for tumour characteristics such as size, histological type, grade, node status, and stage, allow for further assessment of the prognostic changes in the cancers detected [16].
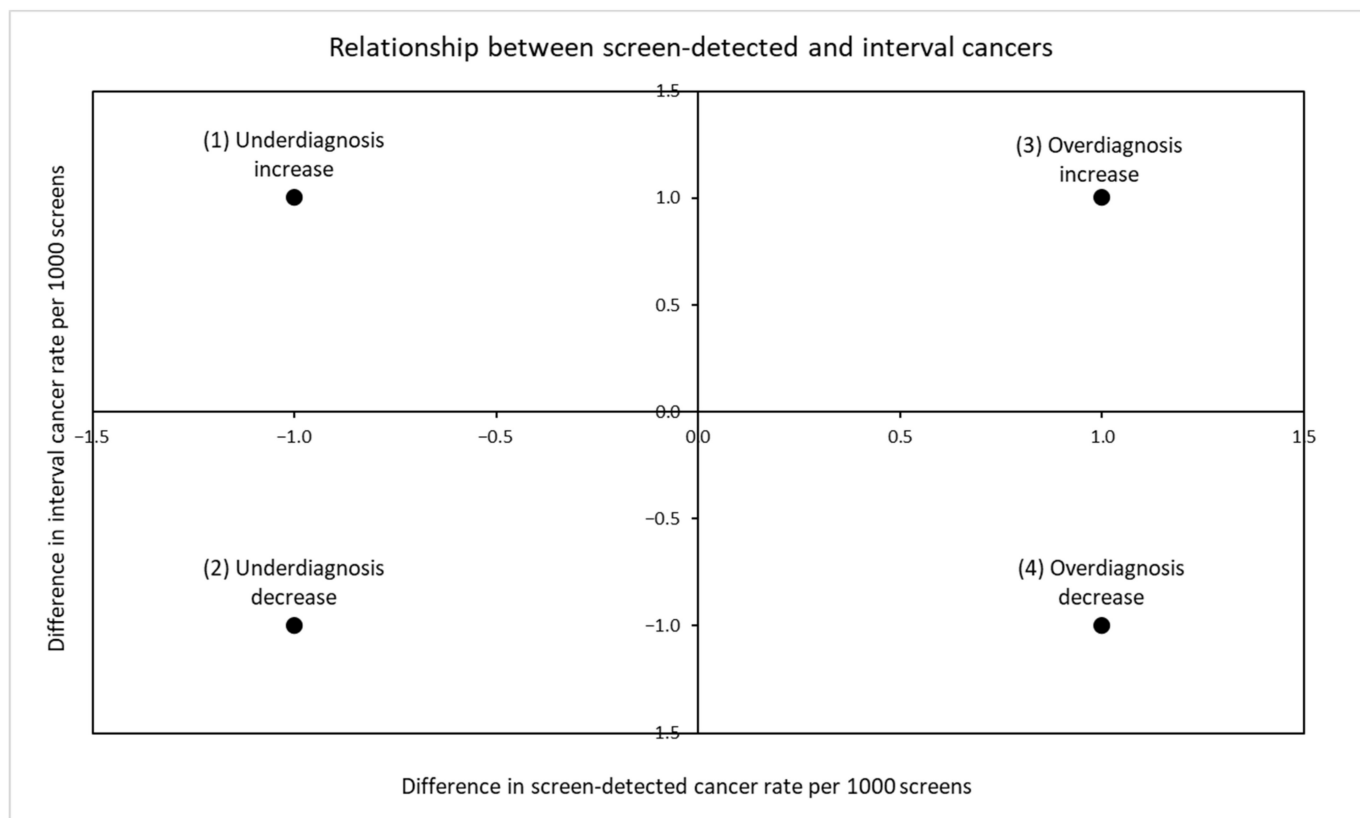
Relationship between screen-detected and interval cancers

Figure 2. Indications from the relationship between screen-detected and interval cancer rates subsequent to a change in a screening program.

4. Previous Uses of Interval Cancers as an Indicator of Benefits and Harms

We conducted a targeted review (R.F. & I.B.) of the literature to examine the way in which interval cancer rates have been used to assess the harms and benefits of changes to screening programs, to identify if similar applied methodologies used interval cancers, as described above. (For search strategy see Figure 3, and for selection of studies see Figure 4) We identified studies that reported on an asymptomatic population at normal risk of the relevant cancer, which reported on either a change to a screening program/technology or to a direct comparison between two or more screening programs/technologies, and assessed both interval cancer rates and screen-detected cancer rates. A summary of the scoping review is presented in Table 1.

#	Search statement
1	interval cancer*.mp.
2	screen*.mp.
3	detect*.mp.

Figure 3. *Cont.*

Int. J. Environ. Res. Public Health **2022**, 19, 14647

5 of 17

4	(sensitivity or specificity).mp. or "Sensitivity and Specificity"/
5	3 or 4
6	compar*.mp.
7	(versus or vs).mp.
8	6 or 7
9	mass screening.mp. or Mass Screening/
10	(screen* adj3 program*).mp.
11	9 or 10
12	1 and 2 and 5 and 8 and 11

Figure 3. Medline Search Strategy. asterisk (*) represents a truncation search.
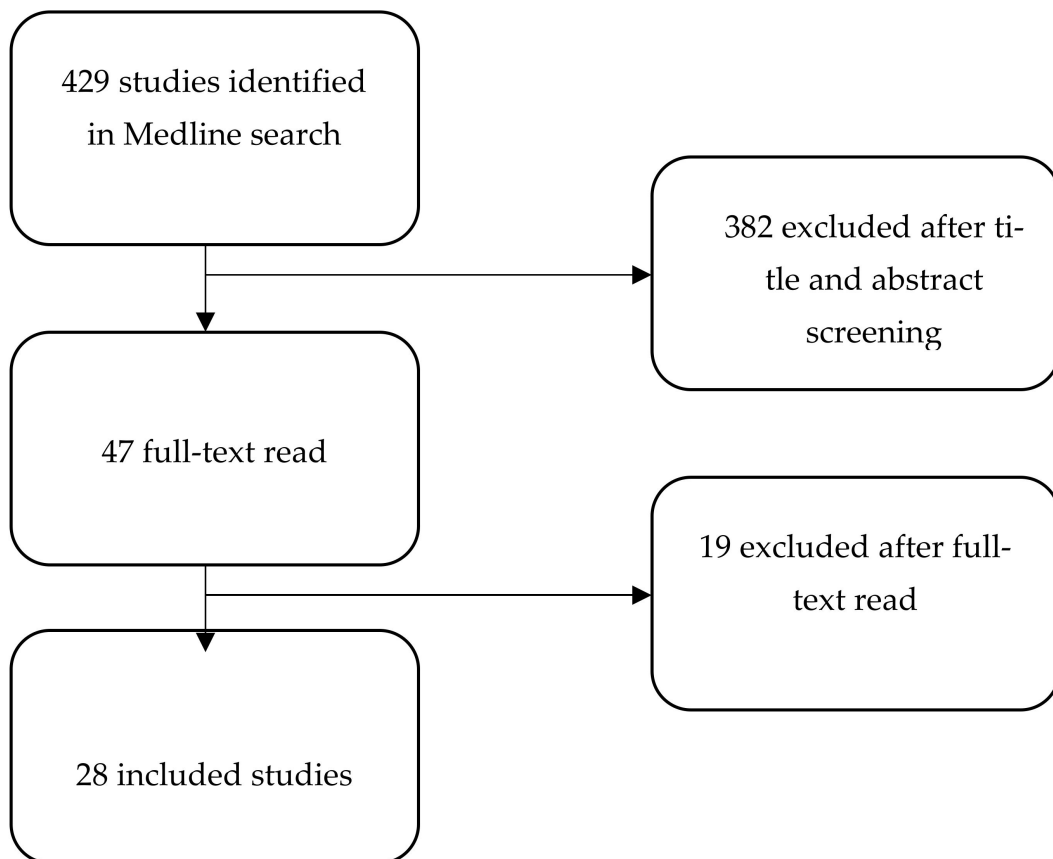

Figure 4. Selection of included studies.

Table 1. Summary of selected studies from scoping review.

Reference	Study Design	Country	Study Time Points	Change/Comparison	Screen-Detected Cancer Rate	Interval Cancer Rate	Harms and/or Benefits
Armaroli (2022) [17]	Randomised control trial	Italy	2014–2020	Digital mammography (DM) versus digital breast tomosynthesis (DBT)	Significantly higher with DBT	Did not differ significantly between the two arms	Compatible either with greater lead time gain with DBT screening in comparison with DM screening and/or with an excess of detection of indolent cancers (overdiagnosis) in the DBT arm.
Bretagne (2021) [18]	Cohort	France	2010–2017	Faecal immunochemical testing (FIT) versus guaiac faecal occult blood testing (gFOBT).	Higher screen-detected ancer rate with FIT	Lower interval cancer rate with FIT	Dramatic decrease in the cumulative incidence rates of interval cancers after switching from gFOBT to FIT. These data provide an additional incentive for countries still using gFOBT to switch to FIT.
Hofvind (2021) [19]	Randomised Control Trial	Norway	2016–2020	DBT versus DM	No significant difference in screen-detected breast cancer among women screened with DBT versus DM	No difference in interval cancer rate with DBT	DBT benefits were not at the cost of worse subsequent round screening outcomes.
Johnson (2021) [20]	Cohort	Sweden	2010–2015	one-view DBT and two-view DM versus two-view DM	Sensitivity was higher for DBT than for DM	one-view DBT and two-view DM was lower than DM	Effect of DBT on interval cancer rate in population screening and that it could translate into additional screening benefit
Bernardi (2020) [21]	Cohort	Italy	2013–2016	DBT versus DM	DBT significantly improved early detection measures	Did not significantly reduce ICR (relative to DM screening),	Not well powered for interval cancers but suggesting that it could add benefit as well as adding over-detection in population BC screening.

Table 1. *Cont.*

Reference	Study Design	Country	Study Time Points	Change/Comparison	Screen-Detected Cancer Rate	Interval Cancer Rate	Harms and/or Benefits
Hovda (2019) [22]	Cohort	Norway	2008–2016	DBT versus DM	Higher for DBT	No difference in rates	Findings show that the excess cancers diagnosed with DBT seem to be less aggressive tumors, indicating potential overdiagnosis; cannot conclude about the clinical implications as the study was not designed to investigate overdiagnosis; further studies are needed; interval breast cancer rate examination is essential when evaluating screening effectiveness
Bahl (2018) [23]	Cohort	US	2009–2011 (DM); 2013–2015 (DBT)	DM versus DBT	No difference in rates; higher proportion of screen-detected cancers were invasive rather than in situ when compared with DM	No difference in rates; MRI-detected interval cancers were more likely to be minimal cancers, compared to symptomatic interval cancers	Complete integration of DBT saw higher rates of screen-detected cancers which were invasive and lower rates of screen-detected in situ cancers; although no change in interval cancer rate between screening technologies, a higher proportion of screen-detected cancers are invasive with DBT.
Houssami (2018) [24]	Cohort	Italy	2011–2012	Tomosynthesis 3D-screening versus standard 2D-mammography alone	Screening sensitivity higher for integrated 2D/3D mammography	Lower for tomosynthesis	Much larger screening studies or pooled analyses are necessary in order to examine interval cancer rates arising; small number of interval cancers precludes analytic evaluation and the study was not planned for interval cancer rate comparison, thus results must be interpreted with caution;

Int. J. Environ. Res. Public Health **2022**, *19*, 14647

8 of 17

Table 1. *Cont.*

Reference	Study Design	Country	Study Time Points	Change/Comparison	Screen-Detected Cancer Rate	Interval Cancer Rate	Harms and/or Benefits
Sankatsing (2018) [25]	Cohort	Netherlands	2004–2011	SFM to DM	Increased screen-detected cancer rate for DM	No difference in rates	There is warranted concern that increase in detection of cancers with DM is indicative of overdiagnosis, however, it was demonstrated that invasive cancers were more likely to be detected
Skaane (2018) [26]	Randomised controlled trial	Norway	2010–2012	DM plus DBT versus DM alone	Higher screen-detected cancer rate for DM+DBT	No difference in rates	Beneficial to distinguish between missed and true interval cancers, which this study did not do; a study with longer follow-up periods and more screening rounds would be required to estimate harms in the form of overdiagnosis; it is important to remember that overdiagnosis does not diminish the benefits of mammography in reducing breast cancer mortality
de Munck (2016) [27]	Cohort	The Netherlands	2004–2010	Screen-film mammography (SFM) versus DM	No difference in rates	No difference in rates	DM can safely be used for screening; DM facilitates easier image transfer and may increase efficiency in hospitals; no indicator of overdiagnosis with shift to DM
McDonald (2016) [28]	Cohort	US	2010–2014	DM versus DBT	Increased screen-detected cancer rate for DBT	Decreased interval cancer rate for DBT	Digital breast tomosynthesis was associated with increased cancer detection and decrease interval cancers; it is unlikely that insignificant cancers were detected at an earlier stage (potential overdiagnosis) as the invasive cancer detection rate remained stable; the primary harm of screening was reduced with decreased false-positive examinations

Reference	Study Design	Country	Study Time Points	Change/Comparison	Screen-Detected Cancer Rate	Interval Cancer Rate	Harms and/or Benefits
Prummel (2016) [29]	Cohort	Toronto	2008–2009	Digital computed radiography (CR) versus direct radiography (DR), versus SFM for breast cancer	Lowest screen-detected cancer rate for CR; screen-detected cancer rates for DR and SFM were similar	Highest interval cancer rate for CR; similar interval cancer rates for DR and SFM	Significantly lower cancer detection for CR mammography suggests that CR screening is missing a large number of breast cancers, in comparison with DR and DFM; programs need to monitor the performance of CR separately and assess its continued use over time
Sverzellati (2016) [30]	Randomised controlled trial	Italy	2000–2011	Comparison of two different strategies of lung cancer screening by low-dose computed tomography (LDCT), namely, annual (LDCT1) or biennial (LDCT2) screen.	No difference in rates	No difference in rates	Biennial screening may be at least as efficient as annual screening in terms of screen detection rates and other outcomes (i.e., sensitivity, specificity); reducing the number of LDCT screens is less harmful to patients and more cost-effective
Klompenhouwer (2015) [10]	Cohort	The Netherlands	2009–2011	Blinded double reading versus non-blinded double reading in a biennial mammography screening population	Higher for blinded double reading	Lower for blinded double reading	Tumour characteristics were examined and were comparable between screening strategies; data suggest that potential overdiagnosis does not differ between blinded and non-blinded double reading; cost-effectiveness harms and benefits still required to evaluate screening holistically; cancer mortality is the best measure of screening effectiveness, but requires long-term follow up

Int. J. Environ. Res. Public Health **2022**, *19*, 14647

10 of 17

Table 1. *Cont.*

Reference	Study Design	Country	Study Time Points	Change/Comparison	Screen-Detected Cancer Rate	Interval Cancer Rate	Harms and/or Benefits
Lehman (2015) [31]	Cohort	US	2003–2009	Digital screening mammography with and without computer-aided detection (CAD)	No difference in rates	No difference in rates	No increased benefit of CAD to women; suggestion that insurers pay more for CAD for no clinically beneficial reason
Sala (2015) [32]	Cohort	Spain	1995–2012	SFM to DM	Increased screen-detected cancer rate for DM	No difference in rates	Increased detection rate with the introduction of DM is in part due to earlier diagnosis, however, examination of cancers showed that invasive cancers were more commonly detected DM; the use of digital technology should not be assumed a threat which increases overdiagnosis in screening
Chiang (2014) [33]	Cohort	Taiwan	2004–2009	Two different FITs for colorectal cancer screening: OC-Sensor versus HM-Jack	Lower screen-detected cancer rate for HM-Jack	Higher interval cancer rate for HM-Jack	Efforts to improve the effectiveness of FITs can be measured using the occurrence of interval cancers; other indicators of performance are needed which assess both short-term and long-term outcomes
Dibden (2014) [34]	Cohort	UK	2003–2005	One-view versus two-view mammography	Higher for two-view mammography	Lower for two-view mammograpy	A reduction in interval cancer rates was associated with two-view mammography; the increase in screen-detected cancer associated with two-view mammography is not likely to be due to overdiagnosis

Reference	Study Design	Country	Study Time Points	Change/Comparison	Screen-Detected Cancer Rate	Interval Cancer Rate	Harms and/or Benefits
Hofvind (2014) [35]	Cohort	Norway	1996–2010	SFM versus DM	No difference in rates	No difference in rates	After initial transitional phase, DM saw lower recall and biopsy rates and thus was associated with less harm; relatively small number of cases means studies generally do not including interval cancer rates; larger population-based studies with sufficient follow-up periods are necessary for screening evaluations;
Choi (2012) [36]	Cross-sectional	Korea	2002–2005	Upper-gastrointestinal series (UGIS) versus endoscopy screening for gastric cancer	Higher screen-detected cancer rate for endoscopy	No difference in interval cancer rate	Endoscopy performed better than UGIS with indicators that it may have a larger impact on gastric cancer mortality, however this was not certain; could not distinguish between false-negative and true interval cancers
Hoff (2012) [37]	Retrospective cohort	Norway	2002–2008	SFM versus DM	No difference in rates	No difference in rates	The transition to DM did not reduce the challenge of missed cancers; characteristics differed between cancers missed at SFM and DM; it is necessary to perform analyses which report on the mammographic features in all missed cancers; study differentiated between true and missed cancers
Seigneurin (2009) [38]	Cohort	France	1994–2004	Two-view versus single-view screening mammography	Higher screen-detected cancer rate for two-view mammography	Lower interval cancer rate for two-view mammography	Findings reveal benefits of performing two-view mammography; decrease in interval cancer rate paralleled by increase in the cancer detection rate;

Int. J. Environ. Res. Public Health **2022**, *19*, 14647

12 of 17

Table 1. *Cont.*

Reference	Study Design	Country	Study Time Points	Change/Comparison	Screen-Detected Cancer Rate	Interval Cancer Rate	Harms and/or Benefits
Roobol (2007) [39]	Randomised controlled trials	The Netherlands and Sweden	1993–1995 (Rotterdam study); 1999–2005 (Gothenburg study)	2-year versus 4-year screening intervals for prostate cancer	Higher for 2-year screening interval	No difference in rates	There is potential overdiagnosis for 2-year prostate cancer screening, due to seeing no difference in the interval cancer rate between interval lengths; the appropriateness of length of screening intervals in cancer screening programs can be assessed using interval cancer rates
Skaane (2007) [40]	Randomised controlled trial	Norway	2000–2001	SFM versus DM	Higher screen-detected cancer rate for DM	Lower interval cancer rate for DM	DM has several potential benefits in mammographic screening, not limited to clinical outcomes; reducing interval cancer rates is crucial in breast cancer screening as it represents the benefits of early detection; further studies are needed to make conclusions about harms and benefits
Wai (2005) [41]	Cohort	British Columbia	Prior to 1997 and after 1997 (exact years not reported)	1-year and 2-year screening intervals for women undergoing screening mammography	Higher screen-detected cancer rate for annual mammography screening	No difference in rates	Other measures needed to assess improvement (i.e., prognostic models showing survival rates)
Hunt (1999) [42]	Retrospective cohort	US	1985–1997	Annual versus biennial screening mammography	Lower for annual screening	Lower for annual screening	Low numbers meant that trends did not reach statistical significance; further research needed to examine harms and benefits
Warren (1997) [43]	Cohort	UK	1989–1991	A comparison of the effectiveness of 28 kV (grid) versus 25 kV (no grid) mammographic techniques for breast screening	Higher screen-detected cancer rate for 28 kV grid	Higher interval cancer rate for 28 kV grid	Non-significant difference in cancer detection rate was offset by an increase in interval cancer rate

Most of the included studies in this review reported on changes to breast screening programs and technologies [10,17,19–29,31,32,34,35,37,38,40–43]. Other cancer screening programs examined included gastric cancer [36], colorectal cancer [18,33], and prostate cancer [39]. While most studies compared two types of screening technology approaches (e.g., film versus digital mammography), some studies utilised screen-detected and interval cancer rates to determine optimal screening interval time-frames (e.g., annual versus two-yearly screening intervals) [39,41].

Studies that reflected on the relationship between screen-detected and interval cancers have often used this relationship as an indicator of benefit, but few have considered it as an indicator of harm [4,5,44]. Of the included studies, 11 consider overdiagnosis, however all but four conclude the results reflect it is not of concern [17,21,22,39]. Some studies utilised the relationship between screen-detected and interval cancers as a measure of effectiveness. For example, McDonald et al. found an increase in screen-detected cancer rates and a subsequent decrease in interval cancer rates for digital breast tomosynthesis compared to digital mammography, and noted this as a surrogate outcome for screening benefit [28]. Other studies found the lack of change in screen-detected or interval cancer rates as an indicator of no additional benefit. Lehman et al. utilised the comparison between screen-detected cancer and interval cancer rates to assess the addition of computer-aided detection (CAD) in digital screening mammography. Considering the increased financial burden on women who opt for mammographic screening with CAD, the authors were able to conclude that CAD offers no increase in clinical benefit [31].

The limitations of sole reliance on interval cancers were noted with the importance of complementing the comparison of interval and screen-detected cancers with tumour characteristics, to ensure increases in detection rate are reflecting clinically important cancers [7,37]. In most studies comparing interval cancers with screen-detected cancers, there are consistent findings indicating that interval cancers have worse prognostic features, such as larger tumour size, higher frequency of node metastases, higher histologic grade, and more advanced disease than screen-detected cancers [6].

One example is a study that used interval rates to consider both benefits and harms, as does our framework, comparing tomosynthesis to digital mammography [21]. This study found that a substantial increase in screen-detected cancers did not lead to a commensurate reduction in interval cancers. The authors noted this finding raises the possibility of overdiagnosis, but commented that their study was not sufficiently powered to measure an effect on interval cancer. Nevertheless, they concluded it seemed reasonable to suspect that some of the increased detection from tomosynthesis is contributing to overdiagnosis in population breast cancer screening [21]. Another study proposed that the lack of translation from detection rate to interval cancer rate may be due to a delay in lead time and/or to overdiagnosis [17].

5. Other Considerations

It is important to consider underlying cancer rates in the population, as they can affect the screen-detected and interval cancer rates. Comparisons of interval cancer should be limited to monitoring within screening services or programs, because comparisons between screening programs and countries is limited by heterogeneity [6]. Changes in screening programs often involve a before/after comparison within a single population, rather than a comparison with a control group, and are thus at higher risk of bias. When comparing interval cancer rates to screen-detected cancer rates to measure a change in a screening program, it is important to consider other factors that may have affected these rates, beyond the change of interest. One way to adjust for these temporal changes that may be confounding the relationship is to not only compare overall rates using methodology such as interrupted time series to limit the attribution to the change of interest. Using an interrupted time series limits the selection bias and confounding due to between-group differences [45–47]. However, there may still be challenges to the comparison, even after accounting for background changes over time. One challenge is how best to account for

Int. J. Environ. Res. Public Health **2022**, *19*, 14647

14 of 17

cancers in people who do not comply with rescreening recommendations. Some people who attended at least one screening have a cancer diagnosed beyond the recommended rescreening time. These cancers, which are neither screen detected nor interval, are not included in the Framework's comparison. Care is needed to ensure that omission of these cancers from analysis does not cause a biased estimate of the difference in interval rates.

6. Learning Curves

The changes in screening programs have been associated with an initial increase in false positive rates [48,49]. This change is probably due to adjustment periods, with aspects like new technologies, additional views, or changes in the number of readers [50]. While the increase may be short lived, the impact on those people who were sent down the clinical pathway of being recalled for further investigation has been shown to be associated with significantly adverse psychological effects in both the short and long term [48,51]. Depending on how drastic the change is, even if an increase in false positives is relatively brief, recall rates may take time to stabilise, and this factor must be considered in relation to the frequency at which a new technology is introduced.

7. Other Benefits and Harms

In addition to the impact of changes in screening programs on health outcomes, it is important to consider the workflow, economic, and environmental cost-effectiveness of such changes. An example of this cost-effect analysis is when a change in technology has significantly shortened the acquisition time but has significantly increased the interpretation time [52]. The initial and ongoing cost impact of a change may be in opposite directions. When this is the case, it is important to determine the duration; the change needs to be sustained to outweigh the upfront costs [53]. These considerations should include both the initial impact of the change, for example the environmental cost of a new screening technology, and the ongoing cost reduction in the digitalisation of screens [54,55].

8. Future Changes

The design and implementation of screening programs around the world will continue to change. It is human nature to think that new technology is better, but the translation to improved health outcomes needs to be evaluated [56]. The measurement of interval cancer rates needs to be prioritised in studies evaluating changes in technology, as they are often a recognised evidence gap [21]. The measurement of interval cancers can be used in screening evaluation, linked to data from cancer registries or obtained through academic trials [57]. Emerging areas in screening, such as artificial intelligence, could also benefit from the application of this framework [58]. Often changes are first introduced in symptomatic clinics; while this information is important in evaluating a change, the probability of disease in a symptomatic person may differ substantially from that of an asymptomatic person [59]. Interval cancers should be considered not only in planned changes to cancer screening, but also in unplanned changes, such as when the COVID-19 pandemic forced a move to telehealth models of healthcare [60]. As new evidence emerges, it is important that decision aids are updated with information relevant to the condition of that screening program [61].

9. Conclusions

This framework puts forward a methodology of utlising intermediate outcomes of screen-detected and interval cancers rates to contextualise and weigh the indicators of benefits and harms of changes to a screening program. This approach is important for the consideration of changes in cancer screening practice and policy accounting for both short- and intermediate-term health outcomes.

Int. J. Environ. Res. Public Health **2022**, *19*, 14647

15 of 17

References

1. Marmot, M.G.; Altman, D.G.; Cameron, D.A.; Dewar, J.A.; Thompson, S.G.; Wilcox, M.; Independent UKPoBCS. The benefits and harms of breast cancer screening: An independent review: A report jointly commissioned by Cancer Research UK and the Department of Health (England) October 2012. *Br. J. Cancer* **2013**, *108*, 2205–2240. [CrossRef] [PubMed]
2. Irwig, L.; Houssami, N.; Armstrong, B.; Glasziou, P. Evaluating new screening tests for breast cancer. *BMJ* **2006**, *332*, 678. [CrossRef] [PubMed]
3. Houssami, N.; Irwig, L.; Ciatto, S. Radiological surveillance of interval breast cancers in screening programmes. *Lancet Oncol.* **2006**, *7*, 259–265. [CrossRef]
4. Houssami, N.; Zackrisson, S.; Blazek, K.; Hunter, K.; Bernardi, D.; Lang, K.; Hofvind, S. Meta-analysis of prospective studies evaluating breast cancer detection and interval cancer rates for digital breast tomosynthesis versus mammography population screening. *Eur. J. Cancer* **2021**, *148*, 14–23. [CrossRef] [PubMed]
5. Lauby-Secretan, B.; Scoccianti, C.; Loomis, D.; Benbrahim-Tallaa, L.; Bouvard, V.; Bianchini, F.; Straif, K. Breast-Cancer Screening—Viewpoint of the IARC Working Group. *N. Engl. J. Med.* **2015**, *372*, 2353–2358. [CrossRef] [PubMed]
6. Houssami, N.; Hunter, K. The epidemiology, radiology and biological characteristics of interval breast cancers in population mammography screening. *NPJ Breast Cancer* **2017**, *3*, 1–13. [CrossRef] [PubMed]
7. Gilliland, F.D.; Joste, N.; Stauber, P.M.; Hunt, W.C.; Rosenberg, R.; Redlich, G.; Key, C.R. Biologic characteristics of interval and screen-detected breast cancers. *J. Natl. Cancer Inst.* **2000**, *92*, 743–749. [CrossRef] [PubMed]
8. Chiarelli, A.M.; Edwards, S.A.; Sheppard, A.J.; Mirea, L.; Chong, N.; Paszat, L.; Shumak, R.S.; O'Malley, F.P. Favourable prognostic factors of subsequent screen-detected breast cancers among women aged 50–69. *Eur J. Cancer Prev.* **2012**, *21*, 499–506. [CrossRef]
9. Cowan, W.K.; Angus, B.; Gray, J.C.; Lunt, L.G.; al-Tamimi, S.R. A study of interval breast cancer within the NHS breast screening programme. *J. Clin. Pathol.* **2000**, *53*, 140–146. [CrossRef]
10. Klompenhouwer, E.G.; Voogd, A.C.; den Heeten, G.J.; Strobbe, L.J.A.; de Haan, A.F.J.; Wauters, C.A.; Broeders, M.J.M.; Duijm, L.E.M. Blinded double reading yields a higher programme sensitivity than non-blinded double reading at digital screening mammography: A prospected population based study in the south of The Netherlands. *Eur. J. Cancer* **2015**, *51*, 391–399. [CrossRef]
11. Brawley, O.W.; Paller, C.J. Overdiagnosis in the Age of Digital Cancer Screening. *JNCI J. Natl. Cancer Inst.* **2020**, *113*, 1–2. [CrossRef] [PubMed]
12. Ropers, F.G.; Barratt, A.; Wilt, T.J.; Nicholls, S.G.; Taylor-Phillips, S.; Kramer, B.S.; Esserman, L.J.; Norris, S.L.; Gibson, L.M.; Harris, R.P.; et al. Health screening needs independent regular re-evaluation. *BMJ* **2021**, *374*, n2049. [CrossRef] [PubMed]
13. Kilpelainen, T.P.; Auvinen, A.; Maattanen, L.; Kujala, P.; Ruutu, M.; Stenman, U.-H.; Tammela, T.L.J. Results of the three rounds of the Finnish Prostate Cancer Screening Trial–the incidence of advanced cancer is decreased by screening. *Int. J. Cancer* **2010**, *127*, 1699–1705. [CrossRef] [PubMed]
14. Houssami, N. Overdiagnosis of breast cancer in population screening: Does it make breast screening worthless? *Cancer Biol. Med.* **2017**, *14*, 1–8. [CrossRef]
15. Clift, A.K.; Dodwell, D.; Lord, S.; Petrou, S.; Brady, S.M.; Collins, G.S.; Hippisley-Cox, J. The current status of risk-stratified breast screening. *Br. J. Cancer* **2022**, *126*, 533–550. [CrossRef] [PubMed]
16. Allgood, P.C.; Duffy, S.W.; Kearins, O.; O'Sullivan, E.; Tappenden, N.; Wallis, M.G.; Lawrence, G. Explaining the difference in prognosis between screen-detected and symptomatic breast cancers. *Br. J. Cancer* **2011**, *104*, 1680–1685. [CrossRef]
17. Armaroli, P.; Frigerio, A.; Correale, L.; Ponti, A.; Artuso, F.; Casella, D.; Falco, P.; Favettini, E.; Fonio, P.; Giordano, L.; et al. A randomised controlled trial of digital breast tomosynthesis vs digital mammography as primary screening tests: Screening results over subsequent episodes of the Proteus Donna study. *Int. J. Cancer* **2022**, *151*, 1778–1790. [CrossRef]

18. Bretagne, J.F.; Carlo, A.; Piette, C.; Rousseau, C.; Cosson, M.; Lievre, A. Significant decrease in interval colorectal cancer incidence after implementing immunochemical testing in a multiple-round guaiac-based screening programme. *Br. J. Cancer* **2021**, *125*, 1494–1502. [CrossRef]

19. Hofvind, S.; Moshina, N.; Holen, A.S.; Danielsen, A.S.; Lee, C.I.; Houssami, N.; Aase, H.S.; Akslen, L.A.; Haldorsen, I.S. Interval and Subsequent Round Breast Cancer in a Randomized Controlled Trial Comparing Digital Breast Tomosynthesis and Digital Mammography Screening. *Radiology* **2021**, *300*, 66–76. [CrossRef]

20. Johnson, K.; Lang, K.; Ikeda, D.M.; Akesson, A.; Andersson, I.; Zackrisson, S. Interval Breast Cancer Rates and Tumor Characteristics in the Prospective Population-based Malmo Breast Tomosynthesis Screening Trial. *Radiology* **2021**, *299*, 559–567. [CrossRef]

21. Bernardi, D.; Gentilini, M.A.; De Nisi, M.; Pellegrini, M.; Fanto, C.; Valentini, M.; Sabatino, V.; Luparia, A.; Houssami, N. Effect of implementing digital breast tomosynthesis (DBT) instead of mammography on population screening outcomes including interval cancer rates: Results of the Trento DBT pilot evaluation. *Breast* **2020**, *50*, 135–140. [CrossRef] [PubMed]

22. Hovda, T.; Brandal, S.H.B.; Sebuodegard, S.; Holen, A.S.; Bjorndal, H.; Skaane, P.; Hofvind, S. Screening outcome for consecutive examinations with digital breast tomosynthesis versus standard digital mammography in a population-based screening program. *Eur. Radiol.* **2019**, *29*, 6991–6999. [CrossRef] [PubMed]

23. Bahl, M.; Gaffney, S.; McCarthy, A.M.; Lowry, K.P.; Dang, P.A.; Lehman, C.D. Breast Cancer Characteristics Associated with 2D Digital Mammography versus Digital Breast Tomosynthesis for Screening-detected and Interval Cancers. *Radiology* **2018**, *287*, 49–57. [CrossRef] [PubMed]

24. Houssami, N.; Bernardi, D.; Caumo, F.; Brunelli, S.; Fanto, C.; Valentini, M.; Romanucci, G.; Gentilini, M.A.; Zorzi, M.; Macaskill, P. Interval breast cancers in the 'screening with tomosynthesis or standard mammography' (STORM) population-based trial. *Breast* **2018**, *38*, 150–153. [CrossRef]

25. Sankatsing, V.D.V.; Fracheboud, J.; de Munck, L.; Broeders, M.J.M.; van Ravesteyn, N.T.; Heijnsdijk, E.A.M.; Verbeek, A.L.M.; Otten, J.D.M.; Pijnappel, R.M.; Siesling, S.; et al. Detection and interval cancer rates during the transition from screen-film to digital mammography in population-based screening. *BMC Cancer* **2018**, *18*, 256. [CrossRef]

26. Skaane, P.; Sebuodegard, S.; Bandos, A.I.; Gur, D.; Osteras, B.H.; Gullien, R.; Hofvind, S. Performance of breast cancer screening using digital breast tomosynthesis: Results from the prospective population-based Oslo Tomosynthesis Screening Trial. *Breast Cancer Res. Treat.* **2018**, *169*, 489–496. [CrossRef]

27. de Munck, L.; de Bock, G.H.; Otter, R.; Reiding, D.; Broeders, M.J.; Willemse, P.H.; Siesling, S. Digital vs. screen-film mammography in population-based breast cancer screening: Performance indicators and tumour characteristics of screen-detected and interval cancers. *Br. J. Cancer* **2016**, *115*, 517–524. [CrossRef]

28. McDonald, E.S.; Oustimov, A.; Weinstein, S.P.; Synnestvedt, M.B.; Schnall, M.; Conant, E.F. Effectiveness of Digital Breast Tomosynthesis Compared With Digital Mammography: Outcomes Analysis From 3 Years of Breast Cancer Screening. *JAMA Oncol.* **2016**, *2*, 737–743. [CrossRef]

29. Prummel, M.V.; Muradali, D.; Shumak, R.; Majpruz, V.; Brown, P.; Jiang, H.; Done, S.J.; Yaffe, M.J.; Chiarelli, A.M. Digital Compared with Screen-Film Mammography: Measures of Diagnostic Accuracy among Women Screened in the Ontario Breast Screening Program. *Radiology* **2016**, *278*, 365–373. [CrossRef]

30. Sverzellati, N.; Silva, M.; Calareso, G.; Galeone, C.; Marchiano, A.; Sestini, S.; Sozzi, G.; Pastorino, U. Low-dose computed tomography for lung cancer screening: Comparison of performance between annual and biennial screen. *Eur. Radiol.* **2016**, *26*, 3821–3829. [CrossRef]

31. Lehman, C.D.; Wellman, R.D.; Buist, D.S.M.; Kerlikowske, K.; Tosteson, A.N.A.; Miglioretti, D.L.; Breast Cancer Surveillance Consortium. Diagnostic Accuracy of Digital Screening Mammography With and Without Computer-Aided Detection. *JAMA Int. Med.* **2015**, *175*, 1828–1837. [CrossRef] [PubMed]

32. Sala, M.; Domingo, L.; Macia, F.; Comas, M.; Buron, A.; Castells, X. Does digital mammography suppose an advance in early diagnosis? Trends in performance indicators 6 years after digitalization. *Eur. Radiol.* **2015**, *25*, 850–859. [CrossRef] [PubMed]

33. Chiang, T.-H.; Chuang, S.-L.; Chen, S.L.-S.; Chiu, H.-M.; Yen, A.M.-F.; Chiu, S.Y.-H.; Fann, J.C.-Y.; Chou, C.-K.; Lee, Y.-C.; Wu, M.-S.; et al. Difference in performance of fecal immunochemical tests with the same hemoglobin cutoff concentration in a nationwide colorectal cancer screening program. *Gastroenterology* **2014**, *147*, 1317–1326. [CrossRef] [PubMed]

34. Dibden, A.; Offman, J.; Parmar, D.; Jenkins, J.; Slater, J.; Binysh, K.; McSorley, J.; Scorfield, S.; Cumming, P.; Liao, X.H.; et al. Reduction in interval cancer rates following the introduction of two-view mammography in the UK breast screening programme. *Br. J. Cancer* **2014**, *110*, 560–564. [CrossRef]

35. Hofvind, S.; Skaane, P.; Elmore, J.G.; Sebuodegard, S.; Hoff, S.R.; Lee, C.I. Mammographic performance in a population-based screening program: Before, during, and after the transition from screen-film to full-field digital mammography. *Radiology* **2014**, *272*, 52–62. [CrossRef]

36. Choi, K.S.; Jun, J.K.; Park, E.-C.; Park, S.; Jung, K.W.; Han, M.A.; Choi, I.J.; Lee, H.-Y. Performance of different gastric cancer screening methods in Korea: A population-based study. *PLoS ONE* **2012**, *7*, e50041. [CrossRef]

37. Hoff, S.R.; Abrahamsen, A.-L.; Samset, J.H.; Vigeland, E.; Klepp, O.; Hofvind, S. Breast cancer: Missed interval and screening-detected cancer at full-field digital mammography and screen-film mammography—Results from a retrospective review. *Radiology* **2012**, *264*, 378–386. [CrossRef]

Int. J. Environ. Res. Public Health **2022**, *19*, 14647

17 of 17

38. Seigneurin, A.; Exbrayat, C.; Labarere, J.; Colonna, M. Comparison of interval breast cancer rates for two-versus single-view screening mammography: A population-based study. *Breast* **2009**, *18*, 284–288. [CrossRef]

39. Roobol, M.J.; Grenabo, A.; Schroder, F.H.; Hugosson, J. Interval cancers in prostate cancer screening: Comparing 2- and 4-year screening intervals in the European Randomized Study of Screening for Prostate Cancer, Gothenburg and Rotterdam. *J. Natl. Cancer Inst.* **2007**, *99*, 1296–1303. [CrossRef]

40. Skaane, P.; Hofvind, S.; Skjennald, A. Randomized trial of screen-film versus full-field digital mammography with soft-copy reading in population-based screening program: Follow-up and final results of Oslo II study. *Radiology* **2007**, *244*, 708–717. [CrossRef]

41. Wai, E.S.; D'Yachkova, Y.; Olivotto, I.A.; Tyldesley, S.; Phillips, N.; Warren, L.J.; Coldman, A.J. Comparison of 1- and 2-year screening intervals for women undergoing screening mammography. *Br. J. Cancer* **2005**, *92*, 961–966. [CrossRef] [PubMed]

42. Hunt, K.A.; Rosen, E.L.; Sickles, E.A. Outcome analysis for women undergoing annual versus biennial screening mammography: A review of 24,211 examinations. *AJR Am. J. Roentgenol.* **1999**, *173*, 285–289. [CrossRef] [PubMed]

43. Warren, R.M.; Duffy, S. A comparison of the effectiveness of 28 kV (grid) versus 25 kV (no grid) mammographic techniques for breast screening. *Br. J. Radiol.* **1997**, *70*, 1022–1027. [CrossRef] [PubMed]

44. Chiarelli, A.M.; Blackmore, K.M.; Mirea, L.; Done, S.J.; Majpruz, V.; Weerasinghe, A.; Rabeneck, L.; Muradali, D. Annual vs Biennial Screening: Diagnostic Accuracy Among Concurrent Cohorts Within the Ontario Breast Screening Program. *J. Natl. Cancer Inst.* **2020**, *112*, 400–409. [CrossRef] [PubMed]

45. Brydon, M.; Kephart, G.; Payne, J.I.; Blake, J. Transitioning to Full Field Digital Mammography in Nova Scotia: Using Interrupted Time Series Methods to Study the Impact of Technology Change on Mammography Volumes. *J. Med. Imaging Radiat. Sci.* **2020**, *8*, 8. [CrossRef] [PubMed]

46. Lopez Bernal, J.; Cummins, S.; Gasparrini, A. The use of controls in interrupted time series studies of public health interventions. *Int. J. Epidemiol.* **2018**, *47*, 2082–2093. [CrossRef]

47. Penfold, R.B.; Zhang, F. Use of Interrupted Time Series Analysis in Evaluating Health Care Quality Improvements. *Acad. Pediatrics* **2013**, *13*, S38–S44. [CrossRef]

48. Brett, J.; Austoker, J. Women who are recalled for further investigation for breast screening: Psychological consequences 3 years after recall and factors affecting re-attendance. *J. Public Health Med.* **2001**, *23*, 292–300. [CrossRef]

49. Farber, R.; Houssami, N.; Wortley, S.; Jacklyn, G.; Marinovich, M.L.; McGeechan, K.; Barratt, A.; Bell, K. Impact of Full-Field Digital Mammography Versus Film-Screen Mammography in Population Screening: A Meta-Analysis. *JNCI J. Natl. Cancer Inst.* **2020**, *113*, 16–26. [CrossRef]

50. Antonio, A.L.; Crespi, C.M. Predictors of interobserver agreement in breast imaging using the Breast Imaging Reporting and Data System. *Breast Cancer Res. Treat.* **2010**, *120*, 539–546. [CrossRef]

51. Bossuyt, P.M.; McCaffery, K. Additional patient outcomes and pathways in evaluations of testing. *Med. Decis Mak.* **2009**, *29*, E30–E38. [CrossRef] [PubMed]

52. Berns, E.A.; Hendrick, R.E.; Solari, M.; Barke, L.; Reddy, D.; Wolfman, J.; Segal, L.; DeLeon, P.; Benjamin, S.; Willis, L. Digital and screen-film mammography: Comparison of image acquisition and interpretation times. *Am. J. Roentgenol.* **2006**, *187*, 38–41. [CrossRef] [PubMed]

53. Hiatt, M.D.; Carr, J.J.; Manning, R.L. The length of time necessary to break even after converting to digital mammography. *J. Telemed Telecare* **2000**, *6*, 222–224. [CrossRef] [PubMed]

54. Comas, M.; Arrospide, A.; Mar, J.; Sala, M.; Vilaprinyo, E.; Hernandez, C.; Cots, F.; Martinez, J.; Castells, X. Budget impact analysis of switching to digital mammography in a population-based breast cancer screening program: A discrete event simulation model. *PLoS ONE* **2014**, *9*, e97459. [CrossRef]

55. Moger, T.A.; Holen, A.; Hanestad, B.; Hofvind, S. Costs and Effects of Implementing Digital Tomosynthesis in a Population-Based Breast Cancer Screening Program: Predictions Using Results from the To-Be Trial in Norway. *Pharm. Open* **2022**, *6*, 495–507. [CrossRef]

56. Bucchi, L.; Ravaioli, A.; Baldacchini, F.; Giuliani, O.; Mancini, S.; Vattiato, R.; Falcini, F.; Giorgi Rossi, P.; Campari, C.; Canuti, D.; et al. Annual mammography at age 45–49 years and biennial mammography at age 50–69 years: Comparing performance measures in an organised screening setting. *Eur. Radiol.* **2019**, *29*, 5517–5527. [CrossRef]

57. Baum, J.K.; Hanna, L.G.; Acharyya, S.; Mahoney, M.C.; Conant, E.F.; Bassett, L.W.; Pisano, E.D. Use of BI-RADS 3-probably benign category in the American College of Radiology Imaging Network Digital Mammographic Imaging Screening Trial. *Radiology* **2011**, *260*, 61–67. [CrossRef]

58. Byng, D.; Strauch, B.; Gnas, L.; Leibig, C.; Stephan, O.; Bunk, S.; Hecht, G. AI-based prevention of interval cancers in a national mammography screening program. *Eur. J. Radiol.* **2022**, *152*, 110321. [CrossRef]

59. Batt, J.; Ainsworth, R.; Rayter, Z.; Nickells, J.; Valencia, A. Sensitivity and missed cancer rate in the symptomatic breast clinic-A retrospective cohort study of 40,323 patients. *Breast J.* **2021**, *27*, 248–251. [CrossRef]

60. Bosch, G.; Posso, M.; Louro, J.; Roman, M.; Porta, M.; Castells, X.; Macia, F. Impact of the COVID-19 pandemic on breast cancer screening indicators in a Spanish population-based program: A cohort study. *Elife* **2022**, *11*, 10. [CrossRef]

61. Barratt, A.; Howard, K.; Irwig, L.; Salkeld, G.; Houssami, N. Model of outcomes of screening mammography: Information to support informed choices. *BMJ Br. Med. J.* **2005**, *330*, 936. [CrossRef] [PubMed]