Conflict of interest

Declarations of interest: none

Abbreviations [1]

[1]DCIS=ductal carcinoma in situ; D1,D2,D3,D4= BIRADS breast density category (2003 edition); M=mammogram; US=ultrasound; pos=lymph node positive; neg=lymph node negative; NR= not reported; US= ultrasound; MRI= magnetic resonance imaging; yrs=years;

Abstract

Background

Supplemental screening with MRI or ultrasound increases cancer detection rate (CDR) in women with standard screening mammography. Whether it also reduces interval cancer rate (ICR) is unclear. This study reviewed the evidence evaluating the effect of supplemental imaging on ICR in women undergoing screening mammography.

Materials and Methods

This systematic review included studies that reported both CDR *and* ICR in women undergoing screening mammography alone compared to those undergoing screening mammography with supplemental imaging.

Results

Five studies (3 randomised trials) were eligible. These reported on 142,153 women undergoing mammography screening alone or mammography with supplemental imaging (3 ultrasound and 2 MRI studies). Two studies included a general screening population and 3 included special populations (young, high genetic risk and/or dense breasts). The incremental CDR for supplemental MRI was 14.2–16.5/1000 screens and for ultrasound was 0–4.4/1000 screens. Effect on ICR was variable but evidence of a reduced ICR was more consistent for studies using supplemental MRI (ICR 0.3–0.8 per 1000 screens) than those using ultrasound (ICR 0.49–1.9 per 1000 screens). The higher CDR and lower ICR with supplemental screening were associated with higher recall and biopsy rates particularly with supplemental MRI (9.5–15.9%, up to 69/1000 screens). Cancers detected with supplemental imaging modalities were generally smaller and earlier stage.

Conclusion

Mammography with supplemental MRI or ultrasound increases detection of cancers (versus mammography only) in some sub-groups but also increases recall and biopsy rates and may have a relatively modest effect in reducing ICR.

Keywords: Breast cancer, Screening, MRI, Supplemental Imaging, Ultrasound, Mammography, Interval Cancer

Introduction

Early detection of breast cancer using mammography screening has been widely implemented as a population-level secondary prevention strategy for breast cancer. Interval breast cancers are not initially detected at mammography screening and emerge clinically before the next screen. Interval cancer rates are monitored in many population screening programs as an indicator of quality and effectiveness of screening.[1] In some population groups, additional or supplemental breast imaging (such as ultrasound or MRI) is used as an adjunct to mammography, particularly in subgroups at relatively higher risk of developing breast cancer (for example, women with a family history of BC or women with dense breasts). This has been shown to increase breast cancer detection in comparison to mammography alone. However, it is unclear whether supplemental screening reduces the interval cancer rate as an effect of detecting more cancers.

This aim of this study was to systematically review the published literature to explore the effect of supplemental screening (MRI or breast ultrasound) compared to mammography alone on cancer detection and interval cancer rates. A further aim was to identify specific groups where supplemental screening is most effective at reducing the rate of interval cancers.

Material and Methods

The review was conducted according to PRISMA recommendations.[2] The primary outcomes of interest were cancer detection rate and interval cancer rate in women undergoing screening mammography, reported in studies that included women who had mammogram screening only and those who had mammogram and supplemental imaging.

Eligibility criteria

Studies were included in the review if they were primary studies that met the following eligibility criteria:

(1) Conducted in a mammography screening or surveillance program (in an asymptomatic population or <10% symptomatic population if study included a mixed symptomatic and asymptomatic population); and

(2) Conducted in a clearly defined population or group, including specifying whether BRCA mutation carriers were included; and

(3) Reporting both screen-detected cancer detection rate *and* interval cancer rate; and

(4) Comparative study design, containing a group undergoing supplemental breast imaging (ultrasound, or MRI) *and* a comparison group undergoing mammography alone.

Exclusion criteria were:

(1) Studies not including a comparison or control group undergoing mammography without a supplemental screening modality.

(2) Case series, multi-reader (observer) studies, or studies of cancer-enriched imaging sets were ineligible as these designs do not allow estimation of primary outcomes. Letters, comments, reviews were ineligible.

Information source and literature search

Medline was searched by one investigator (NH) using the following terms: exploded "Breast neoplasm" combined with "Adjunct$ (adjunct/adjunctive)", *or* ""supplement$" (supplemental)", AND "(mammogra$, or screen$, or surveillance)". The search was restricted to English language and included studies from database inception to August 2020.

Abstracts and full text papers were screened against predefined study eligibility criteria by one author (NH or SY) to determine whether studies met the pre-defined eligibility criteria. Reference lists of eligible studies and relevant reviews of the topic were checked for studies reporting the primary outcomes.

Data collection and data items

Data extraction of eligible studies was performed independently by two authors (SY and NH) and checked by a third author (MB). Discussion and consensus were used to resolve discordance. Evidence tables were constructed containing the following data: authors, year of study, study characteristics (population, methodology), outcomes (interval cancer and cancer detection data/rate; recall or false-positive recall to assessment rate; biopsy rate for needle or surgical biopsy, early (6-month) imaging recommended, measures of test accuracy: sensitivity and specificity, and/or AUC, PPV for recall.)

Risk of bias assessment and data synthesis

Risk of bias of eligible studies was independently assessed by two review authors (SY and EM). This was determined by extracted information of study characteristics and standardised form of the revised Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) checklist[3] as adapted for studies of breast screening by Marinovich et al.[4]

The information collated in the evidence table was used to descriptively summarise the study characteristics, and screening outcomes specifically cancer detection rates and interval cancer rates.

Results

Two hundred and thirty studies were identified in the initial search (*See Figure 1: PRISMA flowchart).*[2] Eleven full-text studies were assessed and five studies[5-9] met all eligibility criteria and were included in the review. These five studies reported data from 142,153 women undergoing breast screening, including 93,172 screened with mammography alone and 56,205 screened with mammography and a supplemental modality (ultrasound n=47,469 or MRI (n= 8,736).

Study characteristics

Study characteristics are summarised in Table 1. Two assessed the role of ultrasound[5,9] and three the role of MRI[6-8] as an adjunct to mammography screening. There were three randomised controlled trials (RCTs)[5,8,9] and two retrospective cohort studies.[6,7] The studies included a range of different populations: one included a population aged less than 55 years at high risk of breast cancer,[9] one included only women with mammographically dense breasts (BIRADS category D),[5] one included only women aged 40–49 years.[8] The remaining two had broad inclusion criteria and included all women of screening age, however only defined subgroups with dense breast tissue or high cancer risk underwent supplemental imaging.[6,7] Women with a personal history of breast cancer were not included in these studies.

Supplemental MRI

There were two studies that assessed the outcomes of screening with supplemental MRI.[5,9]

- Bakker et al[5] was a multicentre randomised control trial (RCT) from the Netherlands. It focussed on women aged 50–75 years, all with extremely dense breast tissue on mammography, randomised to screening with biennial MRI plus mammography (n=4,783) or screening with biennial mammography alone (n=32,312).
- Saadatmand et al[9] was a multicentre RCT from the Netherlands. It focussed on women aged 30–55 years at high risk of breast cancer (lifetime risk ≥20%), randomised to screening with annual MRI with biennial mammography (n= 675) or annual mammography alone (n=680).

Supplemental Ultrasound

There were three studies that assessed the outcomes of screening with supplemental ultrasound.[6-8]

- Lee et al[7] was a retrospective cohort study from USA. It focused on women aged 30–80+ years in a community screening setting. Women were risk matched based on first-degree family history of breast cancer and breast density in a 1:5 ratio, to screening with annual ultrasound plus annual mammography (n=3,386) or annual mammography alone (n=15,176).
- Ohuchi et al[8] was a multicentre RCT from Japan. It focussed on asymptomatic women aged 40-49 years, randomised to screening with annual ultrasound and annual mammography (n=36,859) or annual mammography alone (n=36,139).
- Corsetti et al[6] was a retrospective cohort study from Italy. It focused on women stratified by breast density with dense breast evaluated by annual or biennial mammography and annual or biennial ultrasound (n=7,224) and non-dense breast evaluated by annual or biennial mammography alone (n=12,504).

Study Quality and Risk of bias

The risk of bias, assessed by an adapted QUADAS-2 checklist,[3,4] is shown in the 'traffic light' plot[10] in Figure 2. The risk of bias was low in all studies. No studies that scored 'high risk of bias' for any single item. However, all studies were 'uncertain' for the same criterion of 'application of a reference standard for patients recalled after a positive screening test' as they provided limited details about how registries were used to identify interval cancers. It was also noted that the study from Corsetti assessed outcomes for women with dense compared non-dense breasts (Table 1) raising a possible limitation of the study design[6].

Cancer detection rate (CDR)

Results for CDR are shown in Table 2. The CDR for combined DCIS and invasive cancer ranged from 2.7–7.4 per 1000 screens for mammography alone, 14.2–16.5 for mammography with MRI and incremental detection rate (above mammography) of 5.6 with ultrasound. One study did not report CDR for the mammography alone group.[5] Three studies reported a higher CDR in the supplemental breast imaging group compared to the mammography-alone group.[6,8,9] In these studies, one used MRI[9] and two used ultrasound[6,8] as supplementary modalities. One study found a slightly higher CDR between women screened with mammography alone (7.4) compared to mammography with ultrasound (6.9).[7]

The highest CDR was found in the two studies using MRI which were conducted in very specific populations (CDR 16.5 in women with extremely dense breasts[5] and CDR 14.2 in women with a high familial risk[9]).

Interval cancer rate (ICR)

Results for ICR are also shown in Table 2. ICR for mammography alone ranged from 0.45–5 screens per 1000, and ICR was 0.3–0.8 for MRI as supplemental modality and 0.49–1.9 for ultrasound as supplemental modality.[5-9] Two studies reported lower ICR when supplemental MRI was used,[5,9] although this was not statistically significant in one study.[9] Reported ICRs were variable when supplemental ultrasound was used: one study showed a lower ICR with ultrasound,[8] one showed a lower ICR with mammography alone[7] and one study reported lower ICR when ultrasound was used as a supplemental modality in dense breasts compared to mammography alone in non-dense breasts.[6]

Recall/accuracy/ biopsy rate

Table 3 summarises comparative accuracy including biopsy rate, positive predictive values, sensitivity, and specificity. Higher recall rates were reported for women screened with supplemental imaging compared to those screened with mammography alone.[6-8] The recall rate ranged from 8.8–9% for mammography alone, 15.9% for mammography with MRI and 9.5% for incremental recall rate (above mammography) with MRI. Two studies reported higher recall rates for ultrasound as supplemental modalities than mammography alone. The highest recall rate (12.6%) was reported in the first screening round with mammography with supplemental ultrasound.[7]

Three studies reported higher needle biopsy rates in mammography with supplemental imaging compared to mammography alone screening group.[7-9] The biopsy rate ranged from 17.6-27.7 per 1000 screens (so around 1.8-2.8%) for mammography alone and was much higher at 53–69 per 1000 screens for MRI as supplemental imaging and 45–57 per 1000 screens for US. One study reported no difference in biopsy rates for mammography alone group and US as supplemental modality. One study did not report biopsy rate in mammography alone group.[5] Two studies reported higher PPV for biopsy for mammography alone compared to supplemental imaging with mammography. Surgical biopsy rates were reported in two studies.[6,8] One reported no difference in surgical biopsy rates between the mammography alone group and the group that underwent supplemental US.[8] The other reported additional surgical biopsies in 0.84% of screens for women with dense breasts

undergoing supplemental ultrasound. However, the surgical biopsy rate for women with non-dense breasts (not undergoing supplemental US) was not reported for comparison.[6]

Cancer Characteristics

Table 4 summarises the characteristics of the screen-detected cancers and interval cancers. Three studies reported separate results for screen-detected and interval cancers,[5,6,8] one reported results for screen-detected and interval cancers together[7] and one reported only screen-detected cancer characteristics.[9] For mammography-only screen-detected cancers, the majority were invasive (53%–82%) rather than DCIS (14%–47%). Tumours detected by supplemental screening were also more likely to be invasive: for MRI the proportion of invasive cancers was 81%[5] and 60%[9]; for ultrasound they were 70%[8] and 85%.[6]

For tumours detected by screening with supplemental MRI, median tumour size was consistently smaller (9.0 and 9.5mm) than those detected by mammography alone (17mm in both studies). The median tumour size for tumours detected by screening with supplemental ultrasound did not significantly differ to those detected by mammography alone as shown in the data in Table 4.

Two studies reported that the majority of screen-detected cancers and interval cancers in both mammography only and supplemental imaging (US) did not have node metastases. No data on mortality or long-term outcomes were presented in any of the studies.

The majority of interval cancers in both mammography only and supplemental imaging groups were invasive cancers rather than DCIS.[1,3,9] The proportion of interval cancers that were invasive (vs DCIS) was similar in the two groups: 77–92% invasive (vs 0–23% DCIS) in the mammography only group and 75%–100% (vs 0–25% DCIS) in the supplemental imaging group.[5,6,8] The majority of interval cancers in both groups were <20mm in diameter. There are minimal differences reported in mean tumour size of interval cancers detected by mammography alone (18–20mm) and supplemental imaging (13–20mm).[5,6,8] Despite this, the interval cancers in the supplementary imaging groups were more likely to be early stage than later stage at diagnosis.[5,8]

Discussion

This systematic review evaluated the evidence for breast cancer screening using imaging modalities supplemental to mammography, focusing on studies that reported rates of *both* cancer detection *and* interval cancers. While there were only five studies eligible for inclusion,

the studies were generally of high methodological quality, and they included outcomes for over 142,000 women. However, synthesising the results from these studies was constrained by the heterogeneity in study methodology including supplemental imaging used and the populations included in the source studies. For these reasons, pooling of data for meta-analysis was not appropriate so the results were summarised descriptively. Several findings from this review are relevant to practice and to planning future studies. Firstly, supplemental imaging increased CDR (or led to incremental cancer detection) and this evidence was more consistent for MRI compared to ultrasound. Secondly, there was mixed evidence on the effect that CDR had on the subsequent ICR. Thirdly, the trade-off for improved cancer detection was significant increases in recall and biopsy rates. Each of these issues will be discussed in further detail and in the context of the heterogeneous study populations which might account for the mixed findings between the studies.

Many previous studies have reported increased cancer detection with the addition of supplemental screening for women with normal mammography[11,12] and the results of the present review were consistent with this. The addition of ultrasound to mammography had a modest CDR (up to 7 per 1000 screens) in three studies. The two studies that reported recall rate both found a significant increase in recall rate associated with ultrasound[7,8] (up to 12% or 17 per 1000 screens). For MRI, the CDR was higher (up to 26 per 1000 screens) as was the recall rate (up to 95 per 1000 screens). However, these MRI studies were conducted in young women at high risk of cancer and/or with dense breasts, populations where the baseline mammography CDR is particularly low and yet the underlying breast cancer risk is high. The high recall rate was associated with high screening sensitivity (over 95%) in both MRI studies.[5,9]

The focus of this review was to examine the extent that increased CDR from supplemental screening would impact ICR. Three studies found a reduction in ICR[5,8,9] and one study did not find a reduction in ICR with the use of supplemental screening.[7] The remaining study inferred a reduction in ICR, however the two groups compared had different breast density classification, so there is only a suggestion of a possible reduction in ICR.[6] However, the effect of the reduction in ICR seen across the studies was modest and was much lesser than what would be expect given the magnitude of the increase in CDR. When MRI was used in young, high-risk women (but not known to have gene mutations), the CDR increased from 2.0 to 8.2/1000 screens for invasive cancer, yet the ICR only reduced from 0.6 to 0.3/1000 screens.[9] Similarly, the ultrasound studies found little change in ICR[8] or false negative rate.[7] One

ultrasound study showed a non-significant increase in ICR with supplemental screening, however in this study, only women with dense breasts underwent ultrasound so the underlying risk of breast cancer (and hence risk of an interval cancer) would be expected to be higher.[6] The only study that showed a large reduction in ICR was the MRI study in dense breasts where the ICR reduced from 5.0 to 0.8, where there was also a high incremental CDR of 13.4/1000 for invasive cancer.[5] Here, it should also be noted that this study (Bakker et al) was also the only study that had a biennial (2-yearly) screening interval (whereas the other studies used annual or a mix of annual and biennial screening) as described in Table 1. Given that a larger number of interval cancers occur in the second than the first year in 2-yearly screening practice, it is likely that the combination of participants with dense breasts and the 2-yearly screening interval allowed for a greater MRI effect to be observed on interval cancer rates.[1]

Most interval cancers (detected with and without supplemental imaging) were invasive, node-negative cancers and the tumour size was similar in both groups. However, when supplemental screening was used, the proportion of cancers that were early stage was higher than when mammography alone was used. Only two studies provided this data, so conclusions cannot be drawn regarding differences in interval cancer characteristics.

Both MRI studies showed that cancers detected with MRI screening tended to be smaller than those detected with mammography alone or mammography plus ultrasound and they were also more likely to be node-negative.[5,9] These studies were performed in specific populations: women with extremely dense breasts[5] and young women a high lifetime risk (but not with proven BRCA gene mutations).[9] This may indicate that supplemental MRI may provide some protection from interval cancers that tend to be more aggressive than screen-detected cancers, especially in young women.[13] This also supports the current practice of screening with MRI, with or without mammography, in young women with BRCA gene mutations. While the cancers were generally small and node-negative, long-term data were not presented so the impact of the detection of these cancers on breast cancer mortality is unknown.

This review has strengths and limitations. The strength is the robust review methodology, with strict search criteria, pre-determined study eligibility criteria and the inclusion of a risk of bias score. The limitations are the small number of studies and the heterogeneity of the studies, particularly related to study population and supplemental imaging modality. Importantly, our review explored a targeted question related to supplemental screening and how it affects

cancer detection and the subsequent interval cancer rates, whereas most studies and reviews have only looked at cancer detection. We found evidence suggesting that MRI supplemental screening may be of potential benefit in specific populations at high risk and with dense breasts, whereas the evidence on ultrasound was more limited.

Conclusions

When discussing supplemental screening, women and their physicians must consider the trade-offs. The benefits of increased cancer detection and reduced interval cancer rate are balanced with the potential harms of increased recall, biopsy and cost. This review provides additional information on supplemental screening with MRI or US to aid this discussion. The use of supplemental screening will detect additional cancer but will increase recall and biopsy rate and may only have a modest effect on reducing ICR. Within that context, supplemental screening research and practice may be more appropriately targeted to using MRI as the supplemental modality based on the available evidence. More research is needed to further explore this area.

Funding

Funding: This work was supported by the NBCF Chair in Breast Cancer Prevention program [grant # EC-21-001]; N. Houssami was also supported by National Health and Medical Research Council (NHMRC) Investigator grant [grant #1194410]

References

1. Houssami N, Hunter K. The epidemiology, radiology and biological characteristics of interval breast cancers in population mammography screening. *NPJ Breast Cancer.* 2017;3:12.

2. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n71.

3. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of internal medicine.* 2011;155(8):529-536.

4. Marinovich ML, Hunter KE, Macaskill P, Houssami N. Breast Cancer Screening Using Tomosynthesis or Mammography: A Meta-analysis of Cancer Detection and Recall. *J Natl Cancer Inst.* 2018;110(9):942-949.

5. Bakker MF, de Lange SV, Pijnappel RM, et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. *N Engl J Med.* 2019;381(22):2091-2102.

6. Corsetti V, Houssami N, Ghirardi M, et al. Evidence of the effect of adjunct ultrasound screening in women with mammography-negative dense breasts: interval breast cancers at 1 year follow-up. *Eur J Cancer.* 2011;47(7):1021-1026.

7. Lee JM, Arao RF, Sprague BL, et al. Performance of Screening Ultrasonography as an Adjunct to Screening Mammography in Women Across the Spectrum of Breast Cancer Risk. *JAMA Intern Med.* 2019;179(5):658-667.

8. Ohuchi N, Suzuki A, Sobue T, et al. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): a randomised controlled trial. *Lancet.* 2016;387(10016):341-348.

9. Saadatmand S, Geuzinge HA, Rutgers EJT, et al. MRI versus mammography for breast cancer screening in women with familial risk (FaMRIsc): a multicentre, randomised, controlled trial. *Lancet Oncol.* 2019;20(8):1136-1147.

10. McGuinness LA. robvis: An R package and web application for visualising risk-of-bias assessments. https://github.com/mcguinlu/robvis. Published 2019. Accessed June 2021.

11. Scheel JR, Lee JM, Sprague BL, Lee CI, Lehman CD. Screening ultrasound as an adjunct to mammography in women with mammographically dense breasts. *Am J Obstet Gynecol.* 2015;212(1):9-17.

12. Berg WA, Zhang Z, Lehrer D, et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. *JAMA.* 2012;307(13):1394-1404.

13. Gilliland FD, Joste N, Stauber PM, et al. Biologic characteristics of interval and screen-detected breast cancers. *J Natl Cancer Inst.* 2000;92(9):743-749.

Tables

Table 1: Summary of characteristics of studies comparing mammography alone to mammography with supplemental breast imaging (MRI or ultrasound) on interval cancer rates in population screening

Study author, year (Country) Supplemental modality	Design/ N participants	Study setting/ Participant eligibility	Population characteristics (mean or median age)	Imaging used (and comparison group)	Ascertainment of primary outcomes (with follow-up time)
Bakker, 2019 (Netherlands) MRI	▪ Multicentre randomised controlled trial, 1:4 ratio ▪ MRI+ mammogram (biennial) vs mammogram alone (biennial) ▪ N=40,373 in trial 32,312 mammography alone 8,061 invited for MRI (4,783 accepted MRI)	▪ Dutch population-based digital mammography screening ▪ DENSE trial ▪ Eligibility age 50-75 yrs ▪ All had dense breasts (BIRADS D) on imaging software assessment ▪ Normal mammography at randomisation	▪ Age: 54 years (median) ▪ Genetic risk NR	▪ Intervention: Mammography + MRI 2-yearly ▪ Comparison: Mammography alone 2-yearly	▪ Interval cancer linkage with Netherlands cancer registry ▪ Cancer diagnosed within 24 months after negative results on mammogram included ▪ Follow up MRI 6 months if applicable, next routine mammogram 2 years
Saadatmand, 2019 (Netherlands) MRI	▪ Multicentre randomised controlled trial,1:1 ratio ▪ Mammogram alone (annual) vs MRI (annual)+ mammogram (biennial) ▪ All received annual CBE ▪ N=1355 in trial 680 mammography alone 675 MRI plus mammography	▪ Netherlands outpatient setting ▪ FaMRIsc study ▪ Eligibility age 30-55 yrs ▪ Lifetime breast cancer risk ≥20% all participants (family history without identified mutation or previous history of DCIS) ▪ Previous invasive cancer excluded (DCIS eligible)	▪ Age: 49.4 years (mean) ▪ Density BI-RADS D: 15% ▪ ≥1 first degree relative with breast cancer <50 yrs: 54% MRI group, 58% mammogram group ▪ HRT use never: 88% MRI group; 85% mammogram group	▪ Intervention: Mammography 2-yearly plus MRI yearly ▪ Comparison: Mammography alone yearly (Annual clinical breast examination both groups)	▪ Linkage with Dutch national pathology registry (PALGA) ▪ 12 months or repeat examination at 6 months as per radiological judgement
Lee, 2019 (USA) Ultrasound	▪ Retrospective cohort study ▪ Ultrasound+ mammogram vs mammogram alone ▪ Risk-matched sample (1: 5 ratio) ▪ N=30,062 screens mammogram alone (15,176 women)	▪ USA community screening mammograms using annual screening ▪ Two Breast Cancer Surveillance Consortium registries	▪ Age groups reported, mean/median NR ▪ Supplemental US more likely <50yrs, 1st degree family history, dense breasts ▪ Supplemental US group	▪ Intervention: Mammography + ultrasound ▪ Comparison: Mammography alone	▪ Vermont Breast Cancer Surveillance System and San Francisco Mammography Registry), linked to clinical/ imaging/ pathology databases,

	▪ N=6,081 screens mammogram + US (in 3,386 women) ▪ Screening interval : 95% 1-2 years	▪ Excluded symptomatic and personal history breast cancer ▪ Age range 30 to >80	74% (vs 66%) had dense breasts (Cat C/D) 21% high/very high 5-year risk (vs 14%	▪ (95% had screening interval of 1-2 years)	state tumour registries, SEER programs. ▪ Follow up 12 months or next screening examination
Ohuchi, 2016 (Japan) Ultrasound	▪ Multicentre randomised controlled trial,1:1 ratio ▪ J-START trial ▪ Two screens in two years ▪ N=72,998 in trial 36,139 mammography alone 36,859 mammography plus US ▪ 70% also received annual CBE	▪ Asymptomatic women aged 40–49 years ▪ Previous invasive cancer or DCIS excluded ▪ No inclusion/exclusion criteria re family history	▪ Age: 44 years (mean) ▪ First degree female relatives with breast cancer 0 relatives 95.3% 1 relative 4.6% >1 relative 0.1% ▪ Premenopausal 75·7% Perimenopausal 18·4% Postmenopausal 5.8% ▪ Density NR	▪ Intervention: Mammography + ultrasound (annual) ▪ Comparison: Mammography alone (annual) (with or without clinical breast examination)	▪ Screening records, postal survey or Japan Clinical Research Support Unit checked registries for cancer cases
Corsetti, 2011 (Italy) Ultrasound	▪ Retrospective cohort study ▪ Comparison of interval cancer rates in women with non-dense breasts (mammogram alone) vs dense breasts (mammogram plus US) ▪ To determine interval cancers within 1 year ▪ N= 8865 women, 19,728 screening mammograms 12,504 non-dense 7224 dense	▪ Charity-funded breast service in Italy- screening or for symptom evaluation (symptomatic not included in this study) ▪ Annual or biennial screening; 12-month outcomes measured in this study	▪ Age: 50 years (median) ▪ Subgroups Density BIRADS D1/D2 (N= 8865) ▪ Dense breast D3/&D4 (N=7224)	▪ Intervention: Dense breast group Mammography + ultrasound (annual or biennial) ▪ Comparison: Non-dense breast group Mammography alone (annual or biennial)	▪ 1 year follow-up after screening ▪ Hospital discharge records and cancer registry databases

Table 2: Summary of cancer detection, recall and interval cancer rates for screening with mammography alone compared to mammography with supplemental MRI or ultrasound

Study authors, year (Country) Supplemental modality	Cancer detection rate per 1000 screens (95% CI)		Recall rate (% or per 1000 screens (95% CI))		Interval Cancer rate per 1000 screens (95% CI)	
	Mammography alone	Mammography with Supplemental imaging	Mammography alone	Mammography with Supplemental imaging	Mammography alone	Mammography with Supplemental imaging
Bakker, 2019 (Netherlands) MRI	NR	Incremental to mammography All cancers (with MRI) 16.5 (13.3-20.5) Invasive cancers (with MRI) 13.4 (10.5–17.1)	NR	Incremental to mammography 94.9 (86.9-103.6) per 1000	5.0 (4.3-5.8)	0.8 (CI NR)
Saadatmand, 2019 (Netherlands) MRI	▪ Overall 4.9 (2.6-7.5) ▪ Screen-detected 4·2 (2·0–6·8) ▪ Invasive cancer 2·0 (0·7–3·6) ▪ DCIS 2·3 (0·7–4·3)	▪ Overall 14.2 (10.0-18.8) ▪ Screen-detected 13·9 (9·6–18·5) ▪ Invasive cancer 8·2 (5·0–11·7) ▪ DCIS 5·7 (3·2–8·5)	▪ Overall 276/ 3075 (9.0%) 89.8 per 1000 ▪ Mammography 157/276 (57% of FPs) ▪ CBE 110/276 (40% of FPs)	▪ Overall 449/ 2812 (16.0%) 159.7 per 1000 ▪ Mammography + MRI 19/449 (4% of FPs) ▪ MRI 275/449 (61% of FPs) ▪ Mammography 98/449 (22% of FPs) ▪ CBE 57/449 (13% of FPs)	0.6 (0.0-1.5)	0.3 (0.0-0.9)
Lee, 2019 (USA) Ultrasound	▪ For BIRADS 4,5 or cancer 5.5 (4.7-6.4) ▪ For all cancer 7.4 (6.4-8.4)	▪ For BIRADS 4,5 or cancer 5.4 (3.9-7.6) ▪ For all cancer 6.9 (5.1-9.3)	▪ 9.9 (9.1-10.6) per 1000	▪ 17.6 (17.1-18.0) per 1000	FNR= 1.5 (0.8-2.8)	FNR= 1.9 (1.4-2.4)
Ohuchi, 2016 (Japan)	▪ 3.2	5.0	▪ Recall first screen 3,153/35,965 (8.8%)	▪ Recall first screen 4,647/ 36,752 (12.6%)	▪ 0.97 (CI NR)	0.49 (CI NR)

Ultrasound						
Corsetti, 2011 (Italy) Ultrasound	▪ Non-dense: all=5.3 ▪ Non-dense <50 yrs=2.7 Non-dense ≥50 yrs=6.7	Mammogram-detected ▪ Dense: all=2.8 ▪ Dense <50 yrs=2.7 ▪ Dense ≥50 yrs =2.8 Incremental ultrasound-detected ▪ Dense: all=4.4 ▪ Dense <50 yrs=5.6 ▪ Dense ≥50 yrs =2.8	NR	NR	▪ Non-dense: all=1.0 ▪ Non-dense <50 yrs=0.45 ▪ Non-dense ≥50 yrs=1.	▪ Dense: all=1.1 ▪ Dense <50 yrs=1.5 ▪ Dense ≥50 yrs =0.62

Abbreviations: US= ultrasound; MRI= magnetic resonance imaging; yrs=years; NR= not reported; FNR= False-negative Rate

Table 3: Summary of comparative accuracy of screening with mammography alone compared to mammography with supplemental MRI or ultrasound

Study author, year (Country) Supplemental modality	Biopsy rate (%)		PPV (%)		Sensitivity (%)		Specificity (%)	
	Mammography alone	Mammography with Supplemental imaging	Mammography alone	Mammography with Supplemental imaging	Mammography alone	Mammography with Supplemental imaging	Mammography alone	Mammography with Supplemental imaging
Bakker, 2019 (Netherlands) MRI	NR	▪ 331/4783 (6.9)	NR	All cancer 79/454 (17.4) (BIRADS 3-5) Invasive cancer 64/454 (14.1)	NR	▪ 95.2	NR	▪ 92.0
Saadatmand, 2019 (Netherlands) MRI	▪ 17.6 (per 1000 screens)	▪ 53 (per 1000 screens)	(BIRADS 3-5) ▪ Imaging PPV 4.5 (2.4-7.6) ▪ Biopsy PPV 27.8 (16.5-41.6)	(BIRADS 3-5) ▪ Imaging PPV 8.0 (5.7-10.7) ▪ Biopsy PPV 26.8 (20.0-34.7)	▪ 86·7 (59·5–98·3)	▪ 97·5 (86·8–99·9)	▪ Overall 91·0 (89·9–92·0) ▪ Age <50 yrs 89·6% (88·2–90·9) ▪ Age ≥50 yrs 93·5% (91·9–94·9)	▪ Overall 83·8% (82·4–85·2) ▪ Age <50 yrs 81.9 (80.1-83.6) ▪ Age ≥50 yrs 87·7% (85·4–89·8)
Lee, 2019 (USA) Ultrasound	▪ 27.7 (25.9-29.7)	▪ 57.4 (51.9-63.5)	▪ Biopsy PPV 21.4 (19.6-23.5)	▪ Biopsy PPV 9.5 (6.8-13.1)	▪ 73.8 (68.1- 80.0)	▪ 78.6 (67.1-92.0)	▪ 97.7 (97.6- 97.9)	▪ 94.8 (94.2-95.3)
Ohuchi, 2016 (Japan) Ultrasound	▪ 655/35,695 (1·8%)	▪ 1665/36,752 (4·5%)	NR	NR	▪ 77·0 (70·3–83·7)	▪ 91·1 (87·2–95)	▪ 91·4 (91·1–91·7)	▪ 87·7 (87·3–88·0)
Corsetti, 2011 (Italy) Ultrasound	False positive US 5.5% (395/ 7224) included 61 surgical biopsies	5.5%	NR	NR	▪ Non-dense: all=83.5 ▪ Non-dense <50 yrs=85.7 ▪ Non-dense ≥50 yrs=83.1	▪ Dense: all=86.7 ▪ Dense <50 yrs=80.6 ▪ Dense ≥50 yrs =93.1	NR	NR

	(0.84%) screens with benign outcomes

Table 4: Characteristics of cancers detected by mammography alone compared to cancers detected by mammography with supplemental MRI or ultrasound

		Mammography only		Supplemental Imaging (MRI or US)	
Study author, year (Country) Supplemental modality	**Characteristic**	**Screen detected**	**Interval Cancers**	**Screen Detected**	**Interval Cancers**
Bakker, 2019 (Netherlands) MRI	Number of cancer (n)	NR	161	n=79[1]	4
	Histologic type				
	DCIS	NR	9/161 (5.6%)	NR	0
	Invasive cancers	NR	152/161 (94%)	NR	4/4 (100%)
	Tumour Size (median or mean) OR T-stage[2]	17mm (IQR 12-23)	NR	9.5 mm (IQR 6.8-12)	13.0 mm (IQR 10.5-17.0)
	Lymph Node Status				
	Positive	NR	72/161 (45%)	9/79 (11%)[1]	2/4 (50%)
	Negative	NR	89/161 (55%)	70/79 (87%)[1]	2/4 (50%)
	Anatomic Stage				
	Early (0 or 1)	NR	67/161 (42%)	72/79 (91%)	2/4 (50%)
	Late (II, III, IV)	NR	94/161 (58%)	7/79 (8.9%)	2/4 (50%)
Saadatmand, 2019 (Netherlands) MRI	Number of cancer (n)	15		40	
	Histologic type				
	DCIS	7/15 (47%)	NR	16/40 (40%)	NR
	Invasive cancers	8/15 (53%)	NR	24/40 (60%)	NR
	Tumour Size (median or mean) OR T-stage[2]	Mean 18mm Median 17mm T1a: 0 T1b: 1/8 (13%) T1c: 5/8 (63%) T2: 2/8 (25%) T3:0 T4:0	NR	Mean 12mm Median 9mm T1a:7/24 (29%) T1b: 7/24 (29%) T1c: 7/24 (29%) T2: 2/24 (8%) T3: 1/24 (4%) T4: 0	NR
	Lymph Node Status[2]				
	Positive	5/8 (63%)	NR	4/24 (17%)	NR
	Negative	3/8 (38%)	NR	20/24 (83%)	NR
	Anatomic Stage				
	Early	NR	NR	NR	NR
	Late	NR	NR	NR	NR
Lee, 2019 (USA) Ultrasound	Number of cancer[3] (n)	221		42	
	Histologic type[3]				
	DCIS	77/221 (35%)		20/42 (48%)	
	Invasive cancers	144/221 (65%)		22/42 (52%)	

		Mammography only		Supplemental Imaging (MRI or US)	
Study author, year (Country) Supplemental modality	**Characteristic**	**Screen detected**	**Interval Cancers**	**Screen Detected**	**Interval Cancers**
	Tumour Size (median or mean) OR T-stage [2, 3]	1-5mm: 6% 6-10mm: 26% 11-15mm: 18% 16-20mm: 19% >20mm: 31%		1-5mm: 10% 6-10mm: 25% 11-15mm: 15% 16-20mm: 15% >20mm: 35%	
	Lymph Node Status[2,3]				
	Positive	36/219 (16%)		6/41 (15%)	
	Negative	183/219 (84%)		35/41 (85%)	
	Anatomic Stage[3]				
		(n=219) Stage 0: 35% Stage I: 39% Stage II: 20% Stage III: 6% Stage IV: 1%		(n=41) Stage 0: 49% Stage I: 27% Stage II: 20% Stage III: 5% Stage IV: 0	
Ohuchi, 2016 (Japan) Ultrasound	Number of cancer (n)	117	35	184	18
	Histologic type				
	DCIS	31/117 (27%)	8/35 (23%)	53/184 (29%)	2/18 (11%)
	Invasive cancers	86/117 (74%)	27/35 (77%)	128/184 (70%)	16/18 (89%)
	Tumour Size (mean)	15·1mm (SD8.7)	17·7mm (SD8·0)	15·3mm (SD12·6)	15·3mm (SD8·1)
	Lymph Node Status				
	Positive	34%	37%	18%	38%
	Negative	63%	63%	79%	63%
	Anatomic Stage				
	Early (0-1)	79/117 (68%)	25/35 (71%)	144/184 (78%)	9/18 (50%)
	Late (≥ II)	38/117 (32%)	10/35 (29%)	37/184 (20%)	9/18 (50%)
Corsetti, 2011i (Italy)[4] Ultrasound	Number of cancer (n)	66[5]	13[5]	52[6,7]	8[6,7]
	Histologic type				
	DCIS	9/66 (14%)	0	8/52 (15%)	2/8 (25%)
	Invasive cancers	54/66 (82%)	12/13 (92%)	44/52 (85%)	6/8 (75%)
	Tumour Size (mean or median) OR T-stage	Tis 9/66 (14%) T1 45/66 (68%) T2 7/66 (11%) T3 0 T4 1/66 (1.5%)	Tis 0 T1 12/13 (92%) T2 0 T3 0 T4 0	Tis 8/52 (15%) T1 39/52 (75%) T2/T3/T4 n= 0	Tis 2/8 (25%) T1 6/8 (75%) T2/T3/T4 n= 0
	Lymph Node Status				
	Positive	18/66 (27%)	1/13 (7.7%)	4/52 (7.7%)	0
	Negative	45/66 (68%)	11/13 (85%)	43/52 (83%)	8/8 (100%)
	Anatomic Stage				
	Early				
	Stage 0	9/66 (14%)	0	8/52 (15%)	2/8 (25%)
	Stage 1	31/66 (47%)	12/13 (92%)	35/52 (67%)	6/8 (75%)

		Mammography only		Supplemental Imaging (MRI or US)	
Study author, year (Country) Supplemental modality	**Characteristic**	**Screen detected**	**Interval Cancers**	**Screen Detected**	**Interval Cancers**
	Late (≥II)	23/66 (35%)	0	4/52 (7.7%)	0

Abbreviations: DCIS=ductal carcinoma in situ; D1,D2,D3,D4= BIRADS breast density category (2003 edition); M=mammogram; US=ultrasound; pos=lymph node positive; neg=lymph node negative; NR= not reported

[1] Incremental MRI detected
[2] Invasive Cancer
[3] Cancers combined- screen detected plus interval cancers
[4] Included some cases with unknown tumour stage
[5] Non-Dense Breast
[6] Dense Breast
[7] Mammography + US

Figures

Figure 1: Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) flowchart

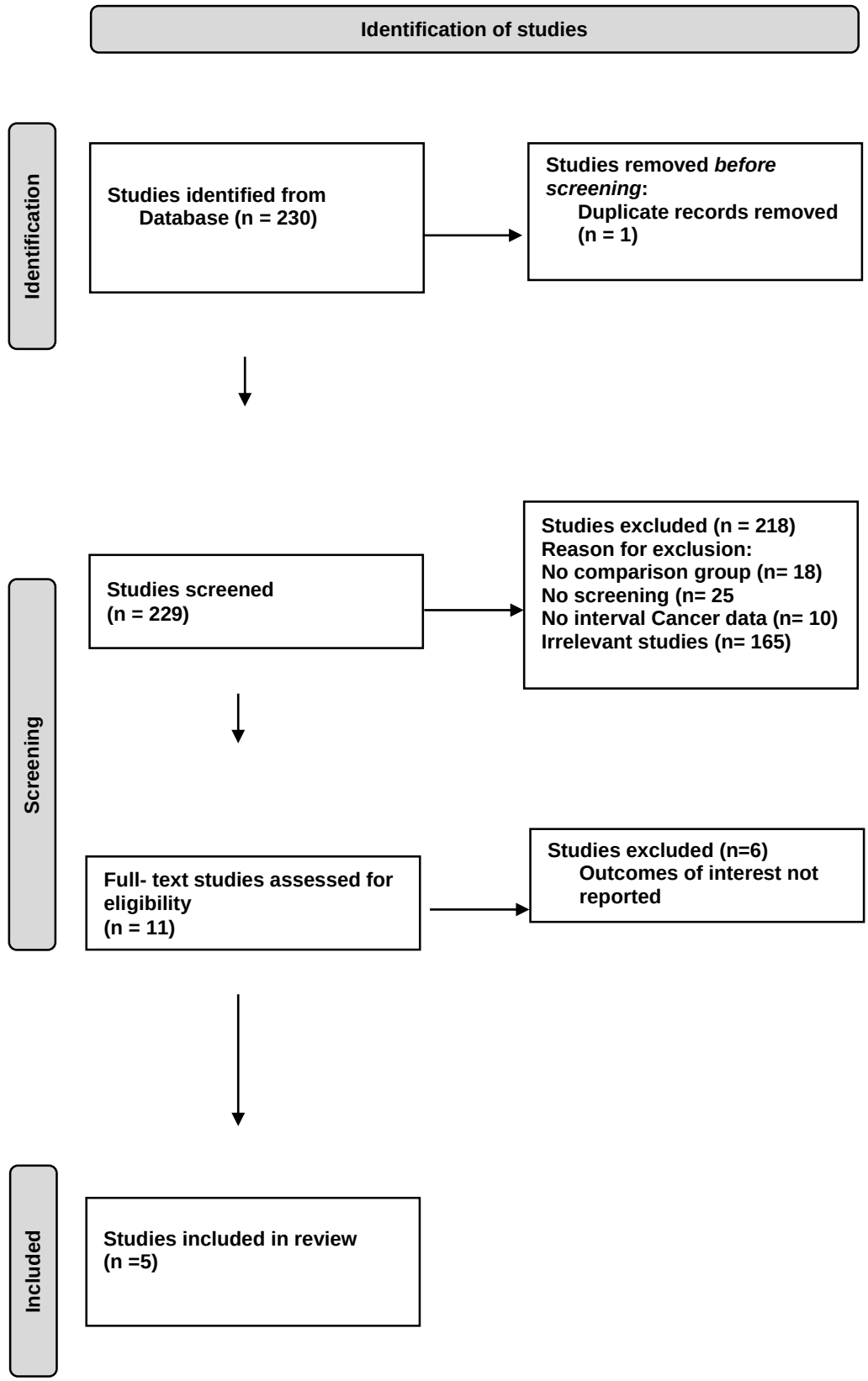
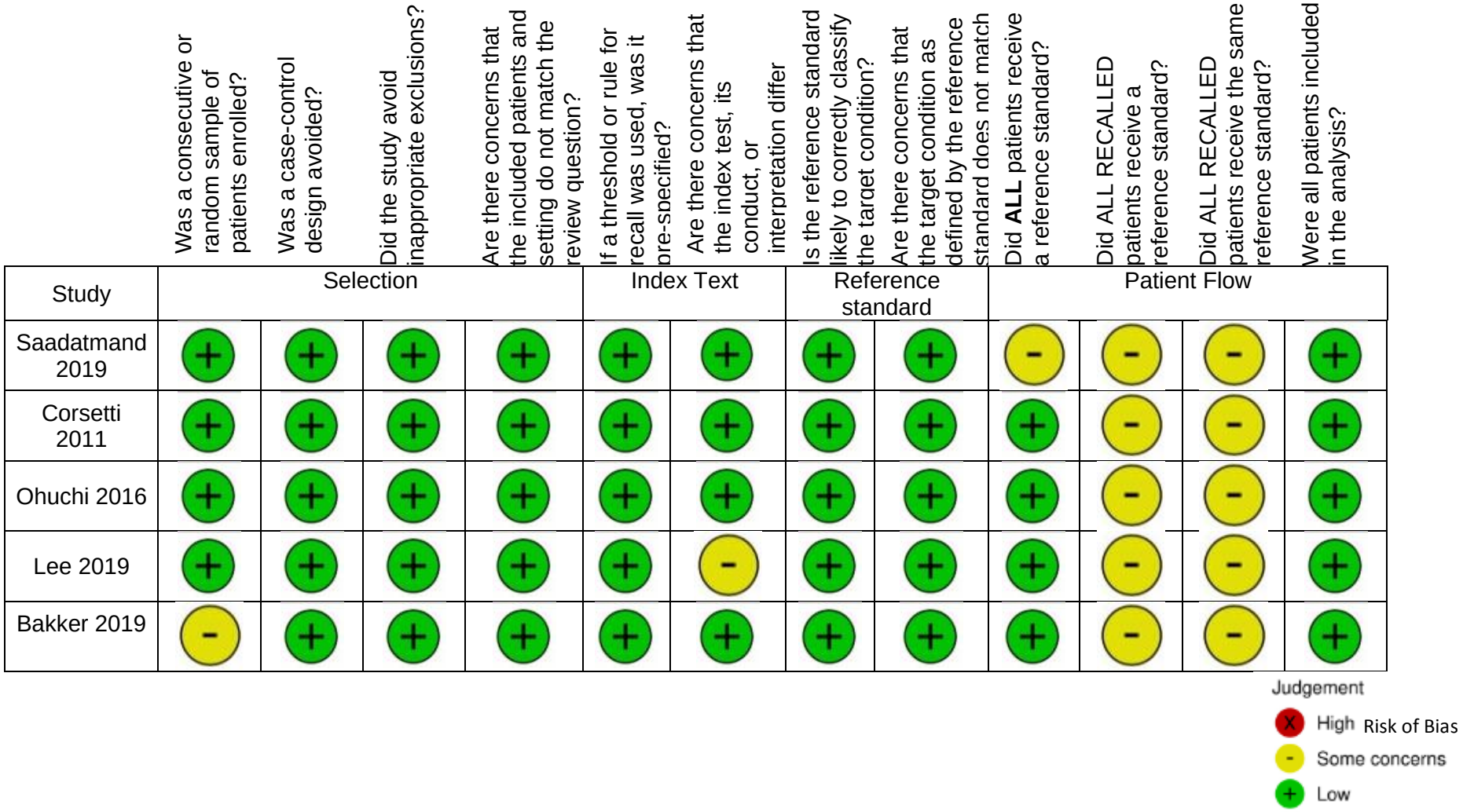


Figure 2: Risk of bias 'traffic light' plot[2]



[2] Risk of bias 'traffic light plot' was created using Robvis[10]