Downloaded from http://jcp.bmj.com/ on January 11, 2017 - Published by group.bmj.com
JCP Online First, published on January 10, 2017 as 10.1136/jclinpath-2016-204184

Estimate of false-positive breast cancer diagnoses from accuracy studies: a systematic review

Nereo Segnan, Silvia Minozzi, Antonio Ponti, Cristina Bellisario, Sara Balduzzi, Marien González-Lorenzo, Silvia Gianola, Paola Armaroli

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/jclinpath-2016-204184).

Department of Cancer Screening, Centre for Epidemiology and Prevention in Oncology (CPO), University Hospital "Città della Salute e della Scienza di Torino", Turin, Italy

Correspondence to
Dr Paola Armaroli, Department of Cancer Screening, Centre for Epidemiology and Prevention in Oncology (CPO), University Hospital "Città della Salute e della Scienza di Torino", Via Cavour 31, Turin 10123, Italy; paola.armaroli@cpo.it

Received 17 October 2016
Revised 12 December 2016
Accepted 13 December 2016

ABSTRACT

Background False-positive histological diagnoses have the same consequences of overdiagnosis in terms of unnecessary treatment. The aim of this systematic review is to assess their frequency at needle core biopsy (CB) and/or surgical excision of the breast.

Methods PubMed, Embase, Cochrane Library were systematically searched up to 30 October 2015. Eligibility criteria: cross-sectional studies assessing diagnostic accuracy of CB compared with surgical excision; studies assessing reproducibility of pathologists reading the same slides. Outcomes: false-positive rates; Misclassification of Benign as Malignant (MBM) histological diagnosis; K statistic. Independent reviewers extracted data and assessed quality using an adapted QUADAS-2 tool.

Results Sixteen studies assessed CB false-positive rates. In 10 studies (41 989 screen-detected lesions), the range of false-positive rates was 0%–7.1%. Twenty-seven studies assessed pathologists' reproducibility. Studies with consecutive, random or stratified samples of all the specimens: at CB the MBM range was 0.25%–2.4% (K values 0.83–0.98); at surgical excision, it was 0.67%–1.2% (K values 0.86–0.94). Studies with enriched samples: the MBM range was 1.4%–6.2% (K values 0.57–0.86). Studies of cases selected for second opinion: the MBM range was 0.29%–12.2% (K values 0.48 and 0.50).

Conclusions High heterogeneity of the included studies precluded formal pooling estimates. When considering studies of higher sample size or methodological quality, false-positive rates and MBM are around 1%. The impact of false-positive histological diagnoses of breast cancer on unnecessary treatment, as well as that of overdiagnosis, is not negligible and is of importance in clinical practice.

INTRODUCTION

Histopathology is currently the main criterion for cancer diagnosis. Consequently, diagnostic pathology errors, that is, misdiagnosis, may lead to incorrect patient management, including delays or unnecessary treatment.[1 2] Pathological diagnosis errors have been shown to exist for nearly every cancer type,[3–5] and the effect of these errors on patient outcome is largely unknown.[2] How to measure and reduce misdiagnosis is a theme of current debate.[6–8]

Implementation of population-based cancer screening has made more critical the issue of false-negative and false-positive diagnoses. In the context of breast cancer (BC), the introduction of mammographic screening and, more recently, of preoperative MRI, has dramatically increased the detection of non-palpable, minimally invasive carcinoma, ductal carcinoma in situ (DCIS) and borderline lesions, which are difficult to diagnose. Published interobserver reproducibility studies have shown low level of agreement among pathologists.[9–13] Non-palpable breast lesions are generally diagnosed by core biopsy (CB). This technique allows to avoid a number of open breast biopsies, but the assessment of CB may be complicated, because a small amount of tissue is obtained. Yet, accurate and reproducible histopathology assessment of breast lesions is of crucial importance when deciding on optimal management choices, which may or may not include treatment.

Moreover, since early diagnosis markedly increases the number of patients with relatively favourable prognosis, concerns about false-positive histological diagnosis and the potential harm of unnecessary treatments for cancer have grown. False-positive histological diagnoses have the same consequences of overdiagnosis (diagnoses of 'cancers' that would not have harmed the patient during lifetime[14]), either in screen-detected or in clinically detected BCs.

In order to assess the risk of false-positive histological diagnoses of BC and the reproducibility of histological diagnoses, we systematically reviewed studies reporting data on the frequency of misclassification of BC in women with suspected malignant lesions undergoing needle CB and/or surgical excision.

METHODS

Data sources and searches

We searched on PubMed, Embase and Cochrane Library all from start date up to 30 October 2015 using a combination of mesh terms and free text words (see online supplementary appendix 1). No language restriction was applied. We inspected the reference lists of the retrieved studies and included relevant articles that we were aware of and that were not captured by our search. Articles suggested by the authors were added to the literature base.

Study selection

Included studies

- Cross-sectional studies assessing the diagnostic accuracy of histological examination of specimen from CB; the CB (index test) was compared with the histological examination of specimen from surgical excision (reference standard) (group 1 studies).
- Studies assessing the agreement among pathologists reading the same CB or surgical excision slides (group 2 studies).

Participants were patients with a suspicion of invasive BC or DCIS, screen and clinically detected. We

To cite: Segnan N, Minozzi S, Ponti A, *et al*. *J Clin Pathol* Published Online First: [*please include* Day Month Year] doi:10.1136/jclinpath-2016-204184

Copyright Article author (or their employer) 2017. Produced by BMJ Publishing Group Ltd under licence.

Downloaded from http://jcp.bmj.com/ on January 11, 2017 - Published by group.bmj.com

considered studies where cases were: (a) all patients consecutively recruited in a certain period of time, randomly selected or stratified samples for type of lesion; (b) enriched samples where a type of cancer or a type of diagnosis (ie, malignant or borderline) were intentionally over-represented or selected to increase the statistical power, when the method for over selection was clearly reported, (c) studies including only cases selected for second opinions. We excluded studies where the criteria and method of case selection were not reported or when the criteria were the selection of 'typical' or 'atypical' or particularly difficult to interpret or 'the most representative' lesion according to the subjective opinion of the study authors.

Outcomes of interest were:

Group 1 studies: the percentage of lesions reclassified from malignant to benign or from invasive cancer to DCIS at CB and surgical excision, respectively.

Group 2 studies: false-positive misclassification of histological diagnosis, the K statistic.

Data extraction and quality assessment

CB, SG, MGL, SM independently selected potentially relevant studies by reading the search hits' titles and abstracts. SM, CB, NS, AP, PA independently assessed for inclusion each full text. CB, SG, MGL, PA extracted data. SM, PA, NS, AP, SB checked the data extraction. Doubts and disagreement were resolved by discussion.

SG, SM, SB independently assessed the methodological quality of the studies. For group 1 studies, we used the QUADAS-2 scale.[15] For group 2 studies, we did not find in the literature any validated checklist specific for this kind of study. Also standards for reporting are lacking in the international literature; a new checklist for quality of reporting was proposed by a group of experts in 2011, not yet validated.[16] So we used a modified checklist adapted from the QUADAS[17] to make it applicable to assess the quality of reliability and agreement studies, in a similar way already done in other systematic reviews.[18]

Data synthesis and analysis

For group 1 studies, we extracted data on the number of patients (or lesions) according to the following definitions:

- CB false-positive rate: that is, percentage of histological diagnoses reclassified from any malignant BC (invasive BC or DCIS) at CB to benign at surgical excision;
  - lesions diagnosed as invasive carcinoma by CB, reclassified as benign tumour by surgical excision;
  - lesions diagnosed as DCIS by CB, reclassified as benign tumour by surgical excision.
- Lesions diagnosed as invasive carcinoma by CB, downgraded to DCIS by surgical excision.

We excluded cases for which it was reported that the lesion found at CB or vacuum-assisted biopsy had been completely excised by the preoperative needle biopsy. Cases receiving neo-adjuvant chemotherapy were also excluded.

For group 2 studies, we extracted data on reproducibility of two (or more) readings of the same specimen calculating histological diagnosis misclassification from available data. When the readers were two, we did not consider one reading as reference unless clearly specified by the authors (ie, expert pathologist). When the readers were more than two, we considered as reference the majority diagnosis, or the consensus diagnosis, or second opinion, or expert diagnosis.

We calculated misclassification as follows:

- Misclassification of Benign lesions as Malignant (MBM): lesions classified as benign at the reference diagnosis and as malignant (DCIS and invasive) at the first diagnosis, on all lesions (or readings, according to available data).
- Misclassification of DCIS as invasive cancers: lesions classified as DCIS at the reference diagnosis and as invasive at first diagnosis, on the total of malignant lesions at reference.
- Misclassification of benign lesions as invasive: lesions classified as benign at the reference diagnosis and as invasive at first diagnosis, on the total of benign lesions at reference.
- Misclassification of benign lesions as DCIS: lesions classified as benign at the reference diagnosis and as DCIS at first diagnosis, on the total of benign lesions at reference,

When the reference was not clearly specified by the authors, we reported the number of discordant diagnoses.

The overall agreement, in terms of Cohen's kappa (K) or as defined by the authors was reported.

We reported results grouping the studies for the following characteristics:

Group 1: studies which included (a) only screen-detected lesions; (b) both screen and clinically detected lesions; (c) type of lesions included not specified.

Group 2: (a) studies with consecutively recruited samples in a certain period of time, randomly selected or stratified samples for type of lesion; (b) studies with enriched samples, where a type of cancer or a type of diagnosis (ie, malignant or borderline) were intentionally over-represented or selected; (c) studies including only cases selected for second opinions.

For each subgroup, the range of the outcomes results was reported. No meta-analyses were performed because of the high heterogeneity of the studies included.

RESULTS

Results of the bibliographic search

After removing duplicates, a total of 6932 records were identified through databases searches. A total 6828 records were excluded on the basis of titles and abstracts as clearly not relevant. One hundred and four articles were judged potentially relevant and acquired in full text. Seven further studies were suggested by the authors as potentially relevant. Sixty-seven articles were excluded (reasons for exclusion in online supplementary appendix 2). Forty-three studies were finally included (44 articles).[9 13 19–59] Dahlstrom *et al* and Sutton *et al*[24 46] reported the results of the same study (figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart). One study was in French, three studies were in Portuguese, all the others in English.

Characteristic of the included studies

In table 1, the number of included studies are reported by type of design, sampling tissue methods and invasion status.

Group 1: studies comparing CB versus surgical excision

Sixteen studies (17 articles) involving a total of 44 713 lesions in women receiving both CB and surgical excision, assessed the false-positive rate of histological examination of specimens from CB compared with surgical excision.[22 24 26 28 29 35 36 38 39 41 46–50 52 54] Ten studies included screen-detected lesions.[22 24 28 29 36 39 41 46–48 50] Three studies included both screen and clinically detected lesions.[26 35 52] Three studies did not report this information.[38 49 54] Recruitment ranged from 1990 to 2012. Detailed study characteristics are reported in online supplementary appendix 3.

Group 2: studies assessing reproducibility of two or more readings of the same specimen

Twenty-seven studies, analysing a total of 13 017 lesions, assessed the reproducibility of two or more readings of the same

Downloaded from http://jcp.bmj.com/ on January 11, 2017 - Published by group.bmj.com

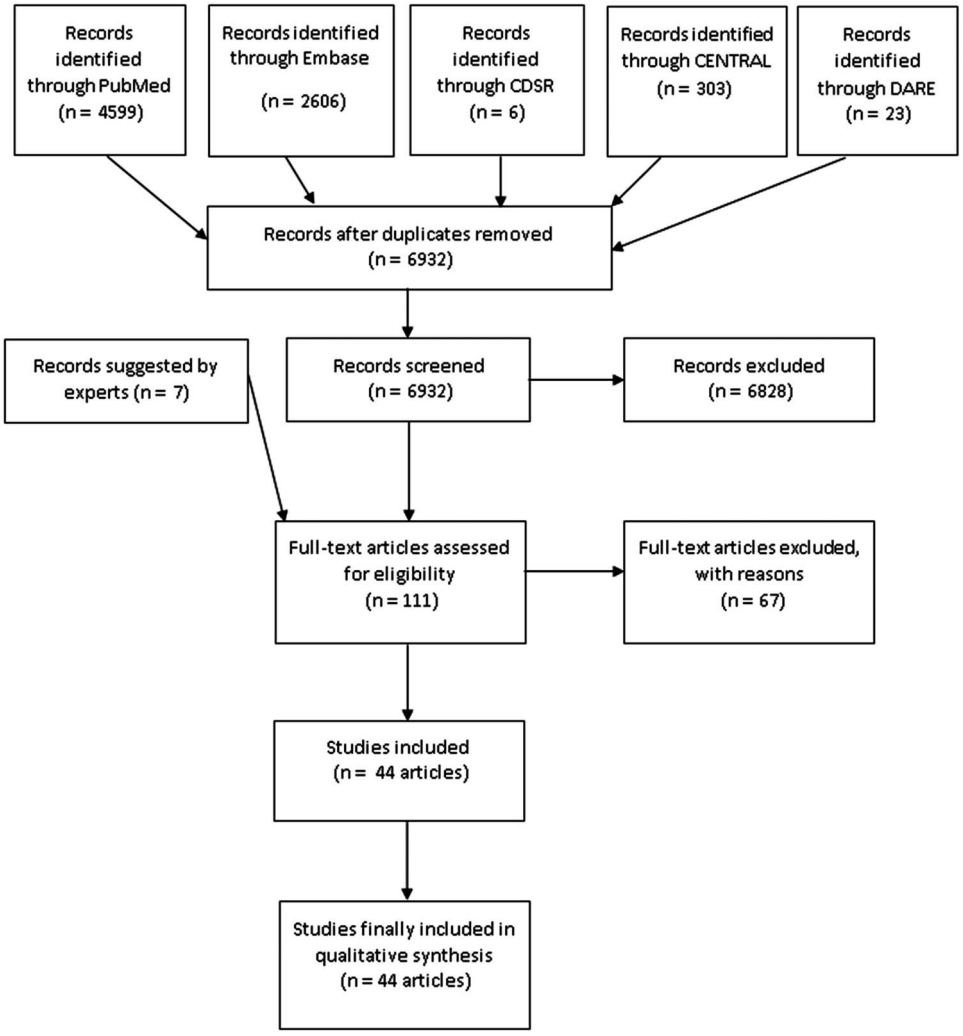

Figure 1 PRISMA flow chart. CENTRAL, Cochrane Central Register of Controlled Trials; CDSR, Cochrane Database of Systematic Reviews; DARE, Database of Abstracts of Reviews of Effects; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Table 1 Number of included studies by type of design, sampling tissue methods and invasion status

		Neoplasia	
Study design	**Tissue sampling**	**DCIS and invasive BC separately**	**Malignant (invasive and DCIS)**
Validation—false-positive rate	CB vs surgical excision (total number=16)	5	16
Reproducibility—misclassification			
Consecutive or random or stratified samples	CB (total number=4)	2	4
	Surgical excision (total number=3)	2	3
	CB and surgical excision (total number=1)	1	1
	Not reported (total number=6*)	2	4
Enriched samples	CB (total number=1)	0	1
	Surgical excision (total number=0)	0	0
	CB and surgical excision (total number=1)	1	1
	Not reported (total number=4)	4	4
Second opinion	CB (total number=0)	0	0
	Surgical excision (total number=2†)	0	1
	CB and surgical excision (total number=4)	4	4
	Not reported (total number=3)	2	3

*One reporting only K statistic.
†From one not possible to calculate outcome from available data.

Downloaded from http://jcp.bmj.com/ on January 11, 2017 - Published by group.bmj.com

Table 2 Core biopsy versus excision at surgery

	1a. Screen-detected lesions						
	Britton *et al*,[22] UK	Dahlstrom *et al*[24] Sutton *et al*,[46] Australia	Jackman *et al*,[28] USA	Lifrange *et al*,[29] Belgium	Rakha *et al*,[36] UK	Seoudi *et al*,[41] USA	Smyth and Cederbom,[39] USA
Diagnosis at CB	15 inadequate (B1) 0 probably benign (B3): 6 probably malignant (B4) 90 malignant (B5)	8 atypical ductal hyperplasia 51 in situ or invasive cancers –	296 benign 135 carcinoma 19 atypical hyperplasia –	Not reported –	9400 B5a (in situ malignancy) 30453 B5b (invasive malignancy) 542 B5c (in situ/invasive malignancy)	16 malignant 8 premalignant –	44 benign 14 malignant –
Diagnosis at surgical excision	10 benign 101 malignant	1 atypical ductal hyperplasia 58 malignant	37 DCIS 45 IDC 26 IDC and EIC 8 ILC	36 benign 64 malignant (2 diagnosed at second surgical biopsy)1 patient refused surgery	40 221 malignant 9 benign 165 true removal of the whole lesion at CB	23 malignant or premalignant 3 normal or benign 1 invasive ductal carcinoma	13 malignant 45 benign
False-positive rate: from malignant BC (invasive BC or DCIS) at CB to benign at SE	Malignant defined as B4+B5: 2.08% (95% CI 0.25% to 7.32%) malignant defined as B5: 0% (97.5% CI 0% to 4.02%)	0% (97.5% CI 0% to 6.98%)	0% (97.5% CI 0% to 3.13%)	3/not reported	0.02% (95% CI 0.01% to 0.04%)	0% (97.5% CI 0% to 20.59%)	7.14% (95% CI 0.18% to 33.87%)

	1a. Screen-detected lesions			1b. Screen and clinically detected lesions		
	Taft 1996,[48] Australia	Vega 1995,[50] Spain	Verkooijen 2002,[47] The Netherlands	Frankel 2011,[26] Brazil	Pijnappel 1997,[35] The Netherlands	Wiratkapun 2010,[52] Thailand
Diagnosis at CB	20 benign 106 malignant (49 invasive ductal cancer, 37 DCIS, 15 invasive lobular cancer, 3 lobular cancer, 2 mixed ductal/lobular)	32 Invasive carcinoma 19 carcinoma 16 atypia 87 benign 2 inadequate cases	352 benign/normal 26 high risk 190 DCIS 290 invasive	88 malignant 1 unsatisfactory	*Palpable lesions:* 22 invasive cancers 5 benign *Non-palpable lesions:* 41 malignancies 35 benign	43 benign 8 high risk 13 malignancy
Diagnosis at surgical excision	108 malignant 18 benign	50 invasive 19 non invasive 87 benign	342 benign/normal 28 high risk 168 DCIS 320 invasive	66 benign 92 malignant	*Palpable lesions:* 5 benign 22 malignant *Non-palpable lesions:* 33 benign 43 malignant	42 benign 5 high risk 17 malignancy
False-positive rate: from malignant BC (invasive BC or DCIS) at CB to benign at surgical excision	0.94%(95% CI 0.02% to 5.14%)	0%(97.5% CI 0% to 6.98%)	1.04%(95% CI 0.35% to 2.41%)	0%(97.5% CI 0% to 4.11%)	Palpable lesions: 0%(97.5% CI 0% to 15.44%)Non-palpable lesions: 0% (97.5% CI 0% to 8.22%)	0%(97.5% CI 0% to 24.71%)

Continued

Downloaded from http://jcp.bmj.com/ on January 11, 2017 - Published by group.bmj.com

Table 2 Continued

	1c. Not specified if clinically or screen-detected lesions		
	Cheurfa 2015,[54] France	Richter-Ehrenstein 2009,[38] Germany	Tse 2010,[49] Hong Kong
Diagnosis at CB	1857 DCIS or invasive cancer	488 malignant (421) Invasive, 67 DCIS 54 B3 lesions, lobular neoplasia or atypical ductal hyperplasia	68 benign 11 atypical 4 suspicious 17 malignant
Diagnosis at surgical excision	(selected only normal/benign after surgery): 26 true removal of the whole lesion at CB · 5 benign	502 malignant (of which 160 DCIS) cases 40 B3 lesions, lobular neoplasia or atypical ductal hyperplasia	63 benign 3 atypical 7 invasive 27 in situ
False-positive rate: from malignant BC (invasive BC or DCIS) at CB to benign at surgical excision	0.27% (95% CI 0.09% to 0.63%)	0.41% (95% CI 0.05% to 1.47%)	0% (97.5% CI 0% to 19.51%)

BC, breast cancer; CB, core biopsy; DCIS, ductal carcinoma in situ.
EIC, extensive intraductal component; IDC , invasive ductal carcinoma; ILC, invasive lobular carcinoma.

specimen among pathologists. Twelve studies included consecutively recruited or randomly selected or stratified samples;[9 13 19 23 25 32 33 43 51 53 55 60] six studies included enriched samples;[20 21 27 42 45 56] nine studies included only cases selected for a second opinion.[30 31 34 37 40 44 57–59]

Types of specimens analysed by type of design and invasion status are reported in table 1.

Two studies assessed the agreement between two readers;[33 43] in 1 study[19] three readers reviewed slides; in 15 studies this information was not reported;[9 23 30 31 34 37 40 44 45 53 55 57–60] in 9 studies, the same set of slides were reviewed by 10,[32] 12,[20 27] 31,[21] 47,[13] 26,[51] 220–466[25] and 186–251 pathologists.[42] Elmore *et al* reported agreement among the initial independent diagnoses of three expert consensus panel members (CPM), and the concordance between 115 pathologists and the consensus-derived reference diagnosis.[56] Recruitment ranged from 1951 to 2014. Detailed study characteristics are reported in online supplementary appendix 3.

Methodological quality of included studies

The overall risk-of-bias assessment for all included studies is reported in online supplementary appendix 4.

Group 1: studies comparing CB versus surgical excision

Twelve of the 16 studies were judged at low, 2 at unclear risk[29 39] and 2 at high[24 46 48] risk of patients selection bias. Concerning applicability, all the studies included patients who matched the review question. Thirteen out of sixteen studies were judged at low risk of bias for the conduct or interpretation of the index test; 2 studies were judged at high risk[24 46 48] and 1 study was judged at unclear risk.[52] Concerning applicability, there was low risk for all the studies that the index test, its conduct or interpretation differed from the review question. Thirteen out of sixteen studies were judged at unclear risk of bias for what concerns the conduct of interpretation of reference standard; Tse *et al*[49] and Verkooijen[47] were judged at low risk of bias, Sutton *et al*, Dahlstrom *et al*[24 46] and Rakha *et al*[36] were judged at high risk of bias. Concerning applicability, there was low concern in all the studies that the target condition as defined by the reference standard does not match the review question. Seven out of sixteen studies were judged at low risk of bias[35 36 39 41 49 50 54] for what concern the patients flow, and 9 at high risk of bias.[22 24 26 28 29 38 46–48 52]

Group 2: studies assessing reproducibility of two or more readings of the same specimen

All but one studies with consecutive or random or stratified samples were judged at low risk of cases selection bias; Palli *et al*[33] was judged at unclear risk for what concern selection of cases because the sample was consecutively recruited for the Swedish series, but it was unclear for the Italian series; all the six studies with enriched samples were judged at high risk of cases selection bias; among the studies with cases selected for a second opinion, six studies were judged at low;[31 37 44 57–59] the other ones at high risk of cases selection bias. Routine clinical data were available for interpretation only in five studies;[13 21 23 31 43] they were not available in three studies;[27 32 33] it was unclear in the other studies. Thirteen out of 27 studies were judged at low risk of bias for what concern independent and blind evaluation of specimens; 3 studies[23 30 31] were judged at high risk and the other ones at unclear risk. All but one study[57] studies were judged at low risk of bias for patients flow.

Downloaded from http://jcp.bmj.com/ on January 11, 2017 - Published by group.bmj.com

Table 2A Core biopsy versus excision at surgery, additional data, available only for screen-detected lesions

Study	From invasive at CB to benign at surgical excision	From DCIS at CB to benign at surgical excision	From invasive at CB to DCIS at surgical excision
Jackman *et al*,[28] USA	0% (97.5% CI 0% to 4.93%)	0% (97.5% CI 0% to 8.22%)	2.74% (95% CI 0.33% to 9.55%)
Lifrange *et al*,[29] Belgium		1/not reported	2/not reported
Rakha *et al*,[36] UK	0.003% (95% CI 0.00008% to 0.018%)	0.08% (95% CI 0.04% to 0.17%)	
Taft *et al*,[48] Australia	0.94% (95% CI 0.02% to 5.14%)		
Verkooijen,[47] The Netherlands	0.34% (95% CI 0.01% to 1.91%)	2.10% (95% CI 0.58% to 5.30%)	1.38% (95% CI 0.38% to 3.49%)

CB, core biopsy; DCIS, ductal carcinoma in situ.

Study results

Group 1: studies comparing CB versus surgical excision

In table 2, results from studies comparing CB versus surgical excision are reported.

When considering studies assessing the false-positive rate of histological examination of screen-detected lesions, the range among 10 studies[22 24 28 29 36 39 41 46–48 50] including 41 989 lesions in total was 0%–7.1%. The false-positive rate was 0% in three studies[26 35 52] comparing histological examination at CB versus surgical excision among 325 total screen and clinically detected lesions, it ranged from 0% to 0.41% among three studies[38 49 54] assessing 2499 lesions whose mode of detection was not reported.

Four studies[22 28 36 48] reported data on screen-detected lesions diagnosed as invasive carcinoma by CB, classified as benign tumour by surgical excision (range 0%–0.94%); three studies[28 36 47] reported data on lesions diagnosed as DCIS by CB, classified as benign tumour by surgical excision (0%–2.1%).

Group 2: studies assessing reproducibility of two or more readings of the same specimen

In table 3, results from studies assessing reproducibility of two or more readings of the same specimen are reported.

Among studies with consecutive or random or stratified samples of all the specimens, MBM of histological diagnosis at CB ranged from 0.25% to 2.4% (four studies), the range of K values was 0.83–0.98 (four studies); at surgical excision MBM ranged from 0.67% to 1.2% (three studies), the K value from 0.86 to 0.94 (three studies); at CB and surgical excision it was estimated as 0.90% (one study), the K value as 0.71; when not reporting the type of specimen, MBM ranged from 0% to 4.8% (four studies), the K value from 0.35 to 0.91 (four studies). In the studies reporting also more detailed data, higher level of misclassification of benign lesion as DCIS is observed (range 0%–3.2%) compared with the levels of misclassification of benign lesions as invasive cancers (range 0%–1.04%). Misclassification of DCIS as invasive cancers, when available, ranged from 0% to 10.8%.

Among studies with enriched samples, MBM of histological diagnosis at CB was 4.4% (one study); at CB and surgical excision it was estimated as 6.2% when reporting data for 115 readers and as 4.03% for the 3 CPM (one study); among studies not reporting the type of specimen, MBM ranged from 1.4% to 3.7% (four studies). The K values range was 0.57–0.86 (five studies). More detailed data were reported by Elmore *et al*:[56] misclassification of benign lesions as invasive cancers at CB and surgical excision was estimated as 0.70% when reporting data for 115 readers and as 0.23% for the 3 CPM; higher values of misclassification of benign lesions as DCIS (9.6% when reporting data for 115 readers and as 6.5% for the 3 CPM) were observed.

Among studies with only cases selected for a second opinion MBM of histological diagnosis at surgical excision was estimated 0.35% (one study); at CB and surgical excision it ranged from 0.93% to 11% (four studies); among studies not reporting the type of specimen MBM ranged from 0.29% to 12.16% (three studies). The K values were 0.48 and 0.50 (two studies).

DISCUSSION

The objective of the review was to assess the frequency of false-positive histological diagnosis of BC in women undergoing CB and/or surgical excision. We included two groups of studies: (1) 16 studies, involving a total of 44 713 lesions, assessing the accuracy and the false-positive rate of histological examination from CB compared with histological examination from surgical excision; (2) 27 studies, analysing a total of 13 017 lesions, assessing the reproducibility of two or more readings of the same specimen among pathologists.

In studies on screen-detected lesions assessing CB accuracy compared with surgical excision as reference standard, the false-positive rate ranged from 0% to 7.1%. In studies with consecutive or random or stratified samples of all the specimens assessing the diagnostic reproducibility among pathologists, the MBM of histological diagnoses ranged from 0.25% to 4.8%. Higher levels of misclassifications were observed for DCIS, and among studies with enriched samples or with cases selected for a second opinion.

The major flaws of studies assessing CB accuracy compared with surgical excision concerned the patients flow and the conduct or interpretation of the reference standard results. The major flaws of studies assessing diagnostic reproducibility among pathologists concerned the method of sample selection and the lack of independent blind evaluation of specimens.

The strength of this review relies in the fact that it provides an overview of all the studies assessing the diagnostic misclassification, in screen and clinically detected BCs retrieved by a comprehensive bibliographic search, in a field which is relatively understudied.[61] On the other hand, a limitation is the high heterogeneity of the included studies, which precluded to formally combine their results in order to give an overall estimate of false-positive histological diagnosis and misclassification rates.

Results from the studies assessing the diagnostic accuracy of CB compared with surgical excision, and from the studies assessing the diagnostic reproducibility among pathologists are different. The majority of the studies comparing CB with surgical excision were conducted with consecutive or representative samples taken from routine practice. If, from this point of view, selection bias is unlikely, on the other side for the majority of them it is unclear whether the pathological diagnosis from surgical excision was blind from or only confirmatory of CB. Also,

Table 3 Reproducibility of two readings of the same specimen

Study	No. of lesion or patients No. of readers No. of total readings	DCIS misclassified as invasive on the total of malignant (DCIS and invasive) lesions (%)	Benign lesions misclassified as invasive on the total of benign lesions (%)	Benign lesions misclassified as DCIS on the total of benign lesions (%)	Benign lesions misclassified as malignant (invasive and DCIS) on the total of readings (or lesions, according to the available data) (MBM) (%)	K value (when specified) or overall agreement or overall disagreement
2a. Patients consecutively recruited or randomly selected samples or stratified samples						
Type of specimen: core needle biopsies						
Screen-detected lesions						
Collins *et al*,[9] USA	No. of lesion or patients: 2004 patients No. of readers: NR No. of total readings: 2004	2.55%	0.12%	0.19%	0.25%	Overall agreement 96.06% (95% CI 95.11% to 96.87%) K overall 0.90 (95% CI 0.88% to 0.92%)
Soofi and Khoury,[60] USA	No. of lesion or patients: 502 lesions No. of readers: NR No. of total readings: 502	NA	NA	NA	2.39%	Overall agreement 79.28% (95% CI 75.47% to 82.74%)
Screen and clinically detected lesions						
Verkooijen *et al*,[13] The Netherlands	No. of lesion or patients: 718 lesions (688 patients) No. of readers: 47 No. of total readings: 718	1.37%	0.36%	3.20%	1.39%	Overall agreement 88.02% (95% CI 85.42% to 90.31%) K overall 0.83 (95% CI 0.78 to 0.88)
Stang *et al*,[43] Germany	No. of lesion or patients: 765 patients No. of readers: 2 No. of total readings: 765	NA	NA	NA	1.96% number of discordant diagnoses	K overall Five-level B-categorisation scheme: 0.89 (95% CI 0.86 to 0.91) Two-level B-categorisation scheme: 0.86 (95% CI: 0.83 to 0.90)
Type of specimen: surgical specimens						
Screen-detected lesions						
Collins *et al*,[9] USA	No. of lesion or patients: 2004 patients No. of readers: NR No. of total readings: 596	1.49%	1.04%	2.59%	1.17%	Overall agreement: 92.62% (95% CI 90.22% to 94.58%) K overall 0.89 (95% CI 0.86 to 0.92)
Screen and clinically detected lesions						
Verkooijen *et al*,[13] The Netherlands	No. of lesion or patients: 718 lesions (688 patients) No. of readers: 47 No. of total readings: 718	1.55%	0%	1.89%	0.69%	Overall agreement 90.39% (95% CI 88.00% to 92.45%) K overall 0.86 (95% CI 0.81 to 0.91)
Not specified if screen or clinically detected lesions						
Middleton *et al*,[55] USA	No. of lesion or patients: 297 lesions No. of readers: NR No. of total readings: 297	NA	NA	NA	0.67%	Overall agreement 86.53% (95% CI 82.11% to 90.20%)
Type of specimen: core needle biopsies and surgical specimens						

Continued

Downloaded from http://jcp.bmj.com/ on January 11, 2017 - Published by group.bmj.com

Table 3 Continued

Study	No. of lesion or patients No. of readers No. of total readings	DCIS misclassified as invasive on the total of malignant (DCIS and invasive) lesions (%)	Benign lesions misclassified as invasive on the total of benign lesions (%)	Benign lesions misclassified as DCIS on the total of benign lesions (%)	Benign lesions misclassified as malignant (invasive and DCIS) on the total of readings (or lesions, according to the available data) (MBM) (%)	K value (when specified) or overall agreement or overall disagreement
Not specified if screen or clinically detected lesions						
Wells *et al*,[51] USA	No. of lesion or patients: 30 lesions No. of readers: 26 No. of total readings: 780	10.84%	0.35%	0.89%	0.90%	K core needle biopsies: 0.98 K surgical specimens: 0.94 K overall: 0.71 Overall agreement 36.67% (95% CI 19.93% to 56.14%)
Type of specimen: not reported						
Screen-detected lesions						
Anderson *et al*,[19] UK	No. of lesion or patients: 875 lesions No. of readers: 3 pathologists reviewed slides in batches of 80–100 cases No. of total readings: 875	NA	NA	NA	0.57%	Overall agreement 94.9% (95% CI 93.27% to 96.30%)K non-invasive/microinvasive=0.84 K non-invasive/invasive=0.91
Not specified if screen or clinically detected lesions						
Chang *et al*,[23] Pennsylvania	No. of lesion or patients: 77 lesions (75 patients) No. of readers: NR No. of total readings: 76	0%	0%	0%	0%	Overall agreement 96.10% (95% CI 89.03% to 99.19%)
Ellis *et al*,[25] UK	No. of lesion or patients: 12 lesions No. of readers: 686 in two circulation No. of total readings: not applicable	NA	NA	NA	NA	K overall 0.78
Palazzo and Hyslop,[32] USA	No. of lesion or patients: 31 lesions No. of readers: 10 No. of total readings: 310	NA	NA	NA	4.84% number of discordant diagnoses	K overall 0.347 K overall among 8 pathologists who used standardised criteria 0.360
Palli *et al*,[33] Sweden and Italy	No. of lesion or patients: 372 lesions No. of readers: 2 No. of total readings: not applicable	4.30% number of discordant diagnoses	NA	NA	NA	Overall agreement 73.92% (95% CI 69.15% to 78.31%) K overall 0.53
Zieger and Stein,[53] Germany	No. of lesion or patients: 1500 lesions (3 groups of 500 consecutive biopsies each) No. of readers: NR No. of total readings: 1500	NA	NA	NA	0.13%	
2b. Enriched samples						
Type of specimen: core needle biopsies						
Screen-detected lesions						

Continued

Downloaded from http://jcp.bmj.com/ on January 11, 2017 - Published by group.bmj.com

Table 3 Continued

Study	No. of lesion or patients No. of readers No. of total readings	DCIS misclassified as invasive on the total of malignant (DCIS and invasive) lesions (%)	Benign lesions misclassified as invasive on the total of benign lesions (%)	Benign lesions misclassified as DCIS on the total of benign lesions (%)	Benign lesions misclassified as malignant (invasive and DCIS) on the total of readings (or lesions, according to the available data) (MBM) (%)	K value (when specified) or overall agreement or overall disagreement
Bianchi *et al*,[21] Italy	No. of lesion or patients: 50 lesions No. of readers: 31 No. of total readings: 1550	NA	NA	NA	4.39% (considering as benign B2 and B3 and malignant B4–B5)	K overall 0.61 (range 0.31–0.88)
Type of specimen: core needle biopsies and surgical specimens						
Not specified if screen or clinically detected lesions						
Elmore *et al*,[56] USA	No. of lesion or patients: 240 lesions (4 test sets of 60 breast biopsies), No. of readers: 115, 3 CPM No. of total readings: 6900; 720	Results from 115 readers: 1.96% Results from 3 CPM: 0.69%	Results from 115 readers: 0.70% Results from 3 CPM: 0.23%	Results from 115 readers: 9.64% Results from 3 CPM: 6.48%	Results from 115 readers: 6.20% Results from 3 CPM: 4.03%	Results from 115 readers: Overall agreement: 75.3%, (95% CI 73.4% to 77.0%) Results from 3 CPM: Unanimous agreement of their independent diagnoses 75% Concordance with the consensus-derived reference diagnoses 90.3%
Type of specimen: not reported						
Screen-detected lesions						
Bianchi *et al*,[20] Italy	No. of lesion or patients: 25 lesions No. of readers: 12 No. of total readings: 300	3.47%	3.85%	0%	2.00%	K overall 0.86 (range 0.65–1.0)
Sloane *et al*,[42] UK	No. of lesion or patients: 12 lesions No. of readers: 186–251 No. of total readings: 17 545	2.31%	1.05%	2.23%	1.36%	K overall Coordinators 0.86 Non-coordinators 0.78
Not specified if screen or clinically detected lesions						
Beck,[45] UK	No. of lesion or patients: 40 lesions No. of readers: 9 No. of total readings: 360	First circulation=0% Second circulation=0%	First circulation=0.89% Second circulation=0%	First circulation=4.44% Second circulation=4.17%	First circulation=3.33% Second circulation=2.5%	K overall 0.57 (value for both circulation and two borderline series combined)
Giardina *et al*,[27] Italy	No. of lesion or patients: 88 lesions No. of readers: 12 No. of total readings: 1032	0%	4.38%	3.19%	3.68%	K overall Between pathologist=0.66 (range 0.57–0.76) Between pathologist and the predominant diagnosis=0.786 (SE 0.27)
2c. Second opinion						
Type of specimen: surgical specimens						
Not specified if screen or clinically detected lesions						
Khazai *et al*,[59] USA	No. of lesion or patients: 1970 lesions No. of readers: NR No. of total readings: 1970	NA	NA	NA	0.35%	_

Continued

Downloaded from http://jcp.bmj.com/ on January 11, 2017 - Published by group.bmj.com

Table 3 Continued

Study	No. of lesion or patients No. of readers No. of total readings	DCIS misclassified as invasive on the total of malignant (DCIS and invasive) lesions (%)	Benign lesions misclassified as invasive on the total of benign lesions (%)	Benign lesions misclassified as DCIS on the total of benign lesions (%)	Benign lesions misclassified as malignant (invasive and DCIS) on the total of readings (or lesions, according to the available data) (MBM) (%)	K value (when specified) or overall agreement or overall disagreement
Renshaw and Gould,[37] USA	No. of lesion or patients: 1131 lesions No. of readers: NR No. of total readings: NA	3/ not reported the number of malignant	NA	NA	NA	-
Type of specimen: core needle biopsies and surgical specimens						
Not specified if screen or clinically detected lesions						
Gomes *et al*,[57] Brazil	No. of lesion or patients: 610 lesions No. of readers: not reported No. of total readings: 610	NA	NA	11.41%	6.23%	
Marco *et al*,[30] Spain	No. of lesion or patients: 205 lesions No. of readers: NR No. of total readings: 205	5.59%	11.76%	0%	0.98%	Overall agreement 74.63% (95% CI 68.00% to 80.44%)
Perez *et al*,[34] Brazil	No. of lesion or patients: 209 lesion No. of readers: NR No. of total readings: 209	6.62%	5.17%	34.48%	11.00%	Overall agreement 83.25% (95% CI 77.49% to 88.05%) K overall 0.5
Romanoff *et al*,[58] USA	No. of lesion or patients: 430 lesions (306 patients) No. of readers: NR No. of total readings: 430	0.74%	NA	2.48%	0.93%	Overall agreement 83.26% (95% CI 79.38% to 86.66%)
Type of specimen: not reported						
Not specified if screen or clinically detected lesions						
Salles *et al*,[40] Brazil	No. of lesion or patients: 329 lesions No. of readers: NR No. of total readings: 329	12.00%	7.75%	23.26%	12.16%	Overall agreement 59.88% (95% CI 54.36% to 65.22%) K overall 0.48
Newman *et al*,[31] Michigan	No. of lesion or patients: 149 patients No. of readers: multidisciplinary tumour board No. of total readings: 149	NA	NA	NA	4.03%	Overall agreement 71.14% (95% CI 63.16% to 78.26%)
Staradub *et al*,[44] USA	No. of lesion or patients: 346 lesions (340 patients) No. of readers: NR No. of total readings: 346	2.03%	NA	NA	0.29%	Overall agreement 80.35% (95% CI 75.76% to 84.40%)

DCIS, ductal carcinoma in situ; NA, not applicable; NR, not reported.

Downloaded from http://jcp.bmj.com/ on January 11, 2017 - Published by group.bmj.com

Downloaded from http://jcp.bmj.com/ on January 11, 2017 - Published by group.bmj.com

the experimental context of the studies does not completely reflect clinical practice, in which, for example, there maybe access to second opinion. Moreover, the majority of these studies report results based on small samples. Among studies assessing screen-detected lesions, only two studies[36 47] have large sample sizes. Yet results of Rakha *et al*[36] are not comparing histological diagnosis at CB and surgical excision in a blind way, assessing false-positive CB diagnosis by collecting further several types of data (ie, radiological ultrasound and clinical opinion, number of CB attempts, number of open surgical procedures to reach a final diagnosis). Taking into account sample size and methodological quality, among the studies assessing the false-positive rate of CB for screen-detected lesions, Verkooijen[47] can be considered an informative study reporting a false-positive rate of 1.04%. A false-positive rate of 0.89% could be estimated adding the results of all the other studies, with a small sample size,[22 24 28 39 41 46 48 50] comparing histological diagnosis at CB and surgical excision for screen-detected lesions. Notwithstanding the high heterogeneity of the design of the studies, a similar overall rough estimate of MBM (0.93%) could be obtained adding the results of all the studies assessing the diagnostic reproducibility among pathologists with consecutive or random or stratified or enriched samples of all the specimens of screen-detected lesions. Considering separately the type of specimen assessed, when adding results of MBM of histological diagnosis at CB only, the estimate is 1.05%; it is 0.87%, when adding results from studies on histological diagnosis at surgical excision only.

To our knowledge, no other systematic reviews with this scope have been published so far. We made every effort to identify all the published literature on this topic by performing sensitive and broad bibliographic searches without any language and date restriction and by contacting authors asking for further information when necessary. Nevertheless, publication bias cannot be ruled out, mainly for these types of studies for which the protocol registration is not required.

Based on the evidence of our review, false-positive histological diagnoses increase unnecessary treatment, in either screen or clinically detected BCs. Although the concordance for invasive BC in the results by Elmore[62] was very good with a discordance of 2.3%, and taking into account that it is very difficult to differentiate between atypia and DCIS, in the same study the variability in the interpretation of individual breast biopsy slides suggest that overinterpretation of the pathological findings may contribute to overtreatment.[8] It has been reported that for women having breast biopsy, nearly one in five (18.5%) with a diagnosis of DCIS would have their biopsy specimen interpreted as atypia or benign by the study's reference consensus panel (with the limiting assumptions that the diagnostic features were present on a single slide and no second opinions were obtained).[62] Moreover in epidemiological studies, a great variation was observed in BC detection rate in population with similar incidence rates. Standardised (European population) total detection rate (×100 000), for women aged 50–69 years, at repeat screening test in 2010 ranges from 2.8 to 5 among the regions in North Italy.[63] Similarly, considerable international variation was found in DCIS detection. The age-standardised detection rate for DCIS varied from 1.6 to 0.45 (per 1000) among 15 screening settings in 12 countries.[64] This variation could not be fully explained by variation in invasive BC incidence or detection rates. One potential cause of such a great difference could be attributed to variability in pathological interpretation. Difficulties in pathological classifications and a poor relationship between morphological features and prognosis cause an appreciable fraction of overdiagnosed cases of BC.[8] Indeed, for example, disagreements in pathological interpretation and classification of intraductal proliferative lesions, as well as in diagnosing microinvasion, could be considered as potential sources of overdiagnosis[8 56] and overtreatment.[65] The false-positive results represent an important harm, with implications for the diagnostic process both for screen and clinically detected lesions. They should be reduced to a minimum, since a diagnosis of a malignant neoplasia implies management decisions and may result in referral to unnecessary and often complex and costly interventions[7] including surgery, radiation or, for false-positive invasive cancers, chemotherapy. In order to better quantify the actual burden of false-positive histological diagnoses, it would be beneficial to conduct studies in which independent reading of slides and consecutive cases or representative samples of cases are analysed. Quality controls of the diagnostic process, including systematic independent reading or second systematic opinion of biopsy samples for borderline lesions of the breast, may reduce the amount of overdiagnosis in asymptomatic populations. Further research is needed to better investigate the implications of these findings for patient management, the causes of the problem and possible remedies.

CONCLUSIONS

When considering studies of higher sample size or methodological quality only, the frequency of histological misclassification of breast benign lesions to malignant lesions in women undergoing core biopsy or surgical excision ranges around 1%, meaning that of 100 positive core biopsies one is false positive and that of 100 surgical open biopsies with any diagnosis one is false positive. The impact of false-positive histological diagnosis on unnecessary treatment, as well as that of overdiagnosis,[66] is not negligible and of importance in clinical practice.

Take home messages

- False positive histological diagnoses have the same consequences of overdiagnosis (diagnoses of "cancers" that would not have harmed the patient during lifetime) either in screen or in clinically detected breast cancers, in terms of overtreatment.
- The aim of this systematic review is to assess false positive histological diagnoses frequency at needle core-biopsy and/or surgical excision.
- When considering studies of higher sample size or methodological quality, false positive rates and misclassification of benign as malignant histological diagnosis are around 1%.
- The impact of false positive histological diagnoses of breast cancer on unnecessary treatment, as well as that of overdiagnosis, is not negligible and is of importance in clinical practice.

Handling editor Cheok Soon Lee

Contributors All the authors contributed in conceptualising the work, performing abstract and full-text review, synthesising the results and drafting and reviewing the manuscript. NS: developed the hypothesis of the study. PA, AP: contributed to study concept. SM: coordinated and revised the review process. CB: performed the literature search. SB: statistical support. SM, CB, PA, MG-L, SG: screened articles for inclusion, abstracted and synthesised data, assessed the methodological quality.

Funding This work was supported by the Centre for Epidemiology and Prevention in Oncology—Piedmont (year 2014, N. 188), National Health Service.

Downloaded from http://jcp.bmj.com/ on January 11, 2017 - Published by group.bmj.com

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Technical appendices are available from the corresponding author at paola.armaroli@cpo.it.

REFERENCES

1 Clary KM, Silverman JF, Liu Y, *et al*. Cytohistologic discrepancies: a means to improve pathology practice and patient outcomes. *Am J Clin Pathol* 2002;117:567–73.
2 Raab SS, Nakhleh RE, Ruby SG. Patient safety in anatomic pathology: measuring discrepancy frequencies and causes. *Arch Pathol Lab Med* 2005;129:459–66.
3 Llewellyn H. Observer variation, dysplasia grading, and HPV typing: a review. *Am J Clin Pathol* 2000;114(Suppl):S21–35.
4 Schlemper RJ, Kato Y, Stolte M. Review of histological classifications of gastrointestinal epithelial neoplasia: differences in diagnosis of early carcinomas between Japanese and Western pathologists. *J Gastroenterol* 2001;36:445–56.
5 Carlson GD, Calvanese CB, Kahane H, *et al*. Accuracy of biopsy Gleason scores from a large uropathology laboratory: use of a diagnostic protocol to minimize observer variability. *Urology* 1998;51:525–9.
6 National Academies of Sciences E, and Medicine. *Improving diagnosis in health care*. Washington, DC: The National Academies Press, 2015.
7 Khullar D, Jha AK, Jena AB. Reducing diagnostic errors–why now? *N Engl J Med* 2015;373:2491–3.
8 Colin C, Devouassoux-Shisheboran M, Sardanelli F. Is breast cancer overdiagnosis also nested in pathologic misclassification? *Radiology* 2014;273:652–5.
9 Collins LC, Connolly JL, Page DL, *et al*. Diagnostic agreement in the evaluation of image-guided breast core needle biopsies: results from a randomized clinical trial. *Am J Surg Pathol* 2004;28:126–31.
10 Elston CW, Sloane JP, Amendoeira I, *et al*. Causes of inconsistency in diagnosing and classifying intraductal proliferations of the breast. European Commission Working Group on Breast Screening Pathology. *Eur J Cancer* 2000;36:1769–72.
11 Palli D, Galli M, Bianchi S, *et al*. Reproducibility of histological diagnosis of breast lesions: results of a panel in Italy. *Eur J Cancer* 1996;32A:603–7.
12 Rosai J. Borderline epithelial lesions of the breast. *Am J Surg Pathol* 1991;15:209–21.
13 Verkooijen HM, Peterse JL, Schipper ME, *et al*. Interobserver variability between general and expert pathologists during the histopathological assessment of large-core needle and open biopsies of non-palpable breast lesions. *Eur J Cancer* 2003;39:2187–91.
14 Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst* 2010;102:605–13.
15 Whiting PF, Rutjes AW, Westwood ME, *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.
16 Kottner J, Audige L, Brorson S, *et al*. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol* 2011;64:96–106.
17 Whiting P, Rutjes AW, Reitsma JB, *et al*. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25.
18 Burr JM, Botello-Pinzon P, Takwoingi Y, *et al*. Surveillance for ocular hypertension: an evidence synthesis and economic evaluation. *Health Technol Assess* 2012;16:1–271, iii-iv.
19 Anderson TJ, Sufi F, Ellis IO, *et al*. Implications of pathologist concordance for breast cancer assessments in mammography screening from age 40 years. *Hum Pathol* 2002;33:365–71.
20 Bianchi S, Palli D, Galli M, *et al*. Reproducibility of histological diagnoses and diagnostic accuracy of non palpable breast lesions. *Pathol Res Pract* 1994;190:69–76.
21 Bianchi S, Caini S, Cattani MG, *et al*. Diagnostic concordance in reporting breast needle core biopsies using the B classification—a panel in Italy. *Pathol Oncol Res* 2009;15:725–32.
22 Britton PD, Flower CD, Freeman AH, *et al*. Changing to core biopsy in an NHS breast screening unit. *Clin Radiol* 1997;52:764–7.
23 Chang JH, Vines E, Bertsch H, *et al*. The impact of a multidisciplinary breast cancer center on recommendations for patient management: the University of Pennsylvania experience. *Cancer* 2001;91:1231–7.
24 Dahlstrom JE, Jain S, Sutton T, *et al*. Diagnostic accuracy of stereotactic core biopsy in a mammographic breast cancer screening programme. *Histopathology* 1996;28:421–7.
25 Ellis IO, Coleman D, Wells C, *et al*. Impact of a national external quality assessment scheme for breast pathology in the UK. *J Clin Pathol* 2006;59:138–45.
26 Frankel PP, Esteves VF, Thuler LC, *et al*. [Diagnostic accuracy of the fine needle aspiration cytology and core needle biopsy as a diagnostic method for breast lesions]. *Rev Bras Ginecol Obstet* 2011;33:139–43.
27 Giardina C, Clemente R, Fusilli S, *et al*. Interobserver reproducibility in diagnosis of breast lesions. A panel in Puglia (Italy). GSBioCaM. *Tumori* 1998;84:354–9.
28 Jackman RJ, Nowels KW, Shepard MJ, *et al*. Stereotaxic large-core needle biopsy of 450 nonpalpable breast lesions with surgical correlation in lesions with cancer or atypical hyperplasia. *Radiology* 1994;193:91–5.
29 Lifrange E, Kridelka F, Colin C. Stereotaxic needle-core biopsy and fine-needle aspiration biopsy in the diagnosis of nonpalpable breast lesions: controversies and future prospects. *Eur J Radiol* 1997;24:39–47.
30 Marco V, Muntal T, Garcia-Hernandez F, *et al*. Changes in breast cancer reports after pathology second opinion. *Breast J* 2014;20:295–301.
31 Newman EA, Guest AB, Helvie MA, *et al*. Changes in surgical management resulting from case review at a breast cancer multidisciplinary tumor board. *Cancer* 2006;107:2346–51.
32 Palazzo JP, Hyslop T. Hyperplastic ductal and lobular lesions and carcinomas in situ of the breast: reproducibility of current diagnostic criteria among community- and academic-based pathologists. *Breast J* 1998;4:230–7.
33 Palli D, Bianchi S, Linell F, *et al*. Histopathologic classification of breast cancer in Sweden and Italy: a comparison between two pathologists. *Tumori* 1992;78:247–9.
34 Perez AA, Balabram D, Salles Mde A, *et al*. [Consultation in breast surgical pathology: interobserver diagnostic variability of atypical intraductal proliferative lesions]. *Rev Bras Ginecol Obstet* 2013;35:164–70.
35 Pijnappel RM, van Dalen A, Borel Rinkes IH, *et al*. The diagnostic accuracy of core biopsy in palpable and non-palpable breast lesions. *Eur J Radiol* 1997;24:120–3.
36 Rakha EA, El-Sayed ME, Reed J, *et al*. Screen-detected breast lesions with malignant needle core biopsy diagnoses and no malignancy identified in subsequent surgical excision specimens (potential false-positive diagnosis). *Eur J Cancer* 2009;45:1162–7.
37 Renshaw AA, Gould EW. Reducing false-negative and false-positive diagnoses in anatomic pathology consultation material. *Arch Pathol Lab Med* 2013;137:1770–3.
38 Richter-Ehrenstein C, Muller S, Noske A, *et al*. Diagnostic accuracy and prognostic value of core biopsy in the management of breast cancer: a series of 542 patients. *Int J Surg Pathol* 2009;17:323–6.
39 Smyth AT, Cederbom GJ. Core biopsy of breast lesions. *J La State Med Soc* 1994;146:499–501.
40 Salles Mde A, Sanches FS, Perez AA, *et al*. [Importance of a second opinion in breast surgical pathology and therapeutic implications]. *Rev Bras Ginecol Obstet* 2008;30:602–8.
41 Seoudi H, Mortier J, Basile R, *et al*. Stereotactic core needle biopsy of nonpalpable breast lesions: initial experience with a promising technique. *Arch Surg* 1998;133:366–72.
42 Sloane JP, Ellman R, Anderson TJ, *et al*. Consistency of histopathological reporting of breast lesions detected by screening: findings of the U.K. National External Quality Assessment (EQA) Scheme. U.K. National Coordinating Group for Breast Screening Pathology. *Eur J Cancer* 1994;30A:1414–19.
43 Stang A, Trocchi P, Ruschke K, *et al*. Factors influencing the agreement on histopathological assessments of breast biopsies among pathologists. *Histopathology* 2011;59:939–49.
44 Staradub VL, Messenger KA, Hao N, *et al*. Changes in breast cancer therapy because of pathology second opinions. *Ann Surg Oncol* 2002;9:982–7.
45 Beck JS. Observer variability in reporting of breast lesions. *J Clin Pathol* 1985;38:1358–65.
46 Sutton S, Dahlstrom JE, Jain S. Stereotactic large-gauge core biopsy: its role in the diagnosis of non-palpable mammographic abnormalities presenting to a screening service. *Australas Radiol* 1997;41:103–8.
47 Verkooijen HM. Diagnostic accuracy of stereotactic large-core needle biopsy for nonpalpable breast disease: results of a multicenter prospective study with 95% surgical confirmation. *Int J Cancer* 2002;99:853–9.
48 Taft R, Chao K, Dear P, *et al*. The role of core biopsy in the diagnosis of mammographically detected lesions. *Aust N Z J Surg* 1996;66:664–7.
49 Tse GM, Tan P-H, Lacambra MD, *et al*. Papillary lesions of the breast—accuracy of core biopsy. *Histopathology* 2010;56:481–8.
50 Vega A, Arrizabalaga R, Garijo F, *et al*. Nonpalpable breast lesion. Stereotaxic core needle aspiration biopsy with a single pass. *Acta Radiol* 1995;36:117–21.
51 Wells WA, Carney PA, Eliassen MS, *et al*. Statewide study of diagnostic agreement in breast pathology. *J Natl Cancer Inst* 1998;90:142–5.
52 Wiratkapun C, Fusuwankaya E, Wibulpholprasert B, *et al*. Diagnostic accuracy of vacuum-assisted stereotactic core needle biopsy for breast lesions. *J Med Assoc Thai* 2010;93:1058–64.
53 Zieger G, Stein H. Retrospective analysis of 1500 mamma biopsies from 1951-1975. A way of self-control in pathology. *Pathol Res Pract* 1982;173:275–82.
54 Cheurfa N, Giard S. [Non-palpable breast cancer malignant on needle core biopsy and no malignancy in surgical excision: how to manage?]. *Gynecol Obstet Fertil* 2015;43:18–24.
55 Middleton LP, Feeley TW, Albright HW, *et al*. Second-opinion pathologic review is a patient safety mechanism that helps reduce error and decrease waste. *J Oncol Pract* 2014;10:275–80.
56 Elmore JG, Longton GM, Carney PA, *et al*. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA* 2015;313:1122–32.

Downloaded from http://jcp.bmj.com/ on January 11, 2017 - Published by group.bmj.com

57 Gomes DS, Porto SS, Balabram D, *et al*. Inter-observer variability between general pathologists and a specialist in breast pathology in the diagnosis of lobular neoplasia, columnar cell lesions, atypical ductal hyperplasia and ductal carcinoma in situ of the breast. *Diagn Pathol* 2014;9:121.
58 Romanoff AM, Cohen A, Schmidt H, *et al*. Breast pathology review: does it make a difference? *Ann Surg Oncol* 2014;21:3504–8.
59 Khazai L, Middleton LP, Goktepe N, *et al*. Breast pathology second review identifies clinically significant discrepancies in over 10% of patients. *J Surg Oncol* 2015;111:192–7.
60 Soofi Y, Khoury T. Inter-institutional pathology consultation: the importance of breast pathology subspecialization in a setting of tertiary cancer center. *Breast J* 2015;21:337–44.
61 Davidson NE, Rimm DL. Expertise vs evidence in assessment of breast biopsies: an atypical science. *JAMA* 2015;313:1109–10.
62 Elmore JG, Nelson HD, Pepe MS, *et al*. Variability in pathologists' interpretations of individual breast biopsy slides: a population perspective. *Ann Intern Med* 2016;164:649–55.
63 Giorgi D, Giordano L, Ventura L, *et al*. [Mammography breast cancer screening in Italy: 2010 survey]. *Epidemiol Prev* 2012;36(Suppl 1):8–27.
64 Lynge E, Ponti A, James T, *et al*. Variation in detection of ductal carcinoma in situ during screening mammography: a survey within the International Cancer Screening Network. *Eur J Cancer* 2014;50:185–92.
65 Ponti A, Lynge E, James T, *et al*. International variation in management of screen-detected ductal carcinoma in situ of the breast. *Eur J Cancer* 2014;50:2695–704.
66 Segnan N, Minozzi S, Armaroli P, *et al*. Epidemiologic evidence of slow growing, nonprogressive or regressive breast cancer: a systematic review. *Int J Cancer* 2016;139:554–73.

Downloaded from http://jcp.bmj.com/ on January 11, 2017 - Published by group.bmj.com

Estimate of false-positive breast cancer diagnoses from accuracy studies: a systematic review

Nereo Segnan, Silvia Minozzi, Antonio Ponti, Cristina Bellisario, Sara Balduzzi, Marien González-Lorenzo, Silvia Gianola and Paola Armaroli

J Clin Pathol published online January 10, 2017

Updated information and services can be found at:
http://jcp.bmj.com/content/early/2017/01/10/jclinpath-2016-204184

These include:

Supplementary Material Supplementary material can be found at:
http://jcp.bmj.com/content/suppl/2017/01/10/jclinpath-2016-204184.DC1.html

References This article cites 65 articles, 8 of which you can access for free at:
http://jcp.bmj.com/content/early/2017/01/10/jclinpath-2016-204184#BIBL

Email alerting service Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/