Centro di Riferimento per l'Epidemiologia
e la Prevenzione Oncologica in Piemonte

www.cpo.it

# Errore diagnostico: revisione della letteratura e motivi di attualità

Workshop 2015
"Programma regionale di screening mammografico Prevenzione Serena"
Torino 2 dicembre 2015

# Stima dei tumori della mammella falsi positivi da studi di accuratezza diagnostica: una revisione sistematica

Nereo Segnan, Silvia Minozzi, Antonio Ponti, Cristina Bellisario,
Sara Balduzzi, Marien González-Lorenzo, Silvia Gianola, Paola Armaroli
SCDO Epidemiologia dei tumori, CPO Piemonte, A.O.U. Città della Salute e della Scienza di Torino

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

## *Background-I*

- L'**istopatologia** è attualmente il principale **criterio** per la **diagnosi di cancro**.
- Gli **errori di patologia diagnostica** posso condurre ad un management scorretto del paziente, inclusi **ritardi nel trattamento** o **trattamenti non necessari**.

## *Background-II*

- L'introduzione dello screening mammografico ha aumentato l'identificazione di carcinomi non palpabili, minimamente invasivi, di DCIS e di lesioni borderline, **difficili da diagnosticare e con prognosi favorevole**.
- Le **diagnosi istologiche false positive aumentano la sovradiagnosi** (diagnosi di 'cancri' che non avrebbero danneggiato il paziente nel corso della vita) nello screening per cancro della mammella ed il **sovratrattamento,** nelle lesioni screen detected ed in quelle diagnosticate clinicamente.

## *Obiettivi*

Valutare la frequenza di diagnosi istologiche false positive di tumore della mammella in donne con sospetto di lesione maligna che effettuano una core needle biopsy (CB) e/o una escissione chirurgica (EC)

## *Materiali e Metodi*

Revisione sistematica di studi che valutano:

-l'accuratezza diagnostica dell'esame istologico da CB confrontato con quello del pezzo operatorio,

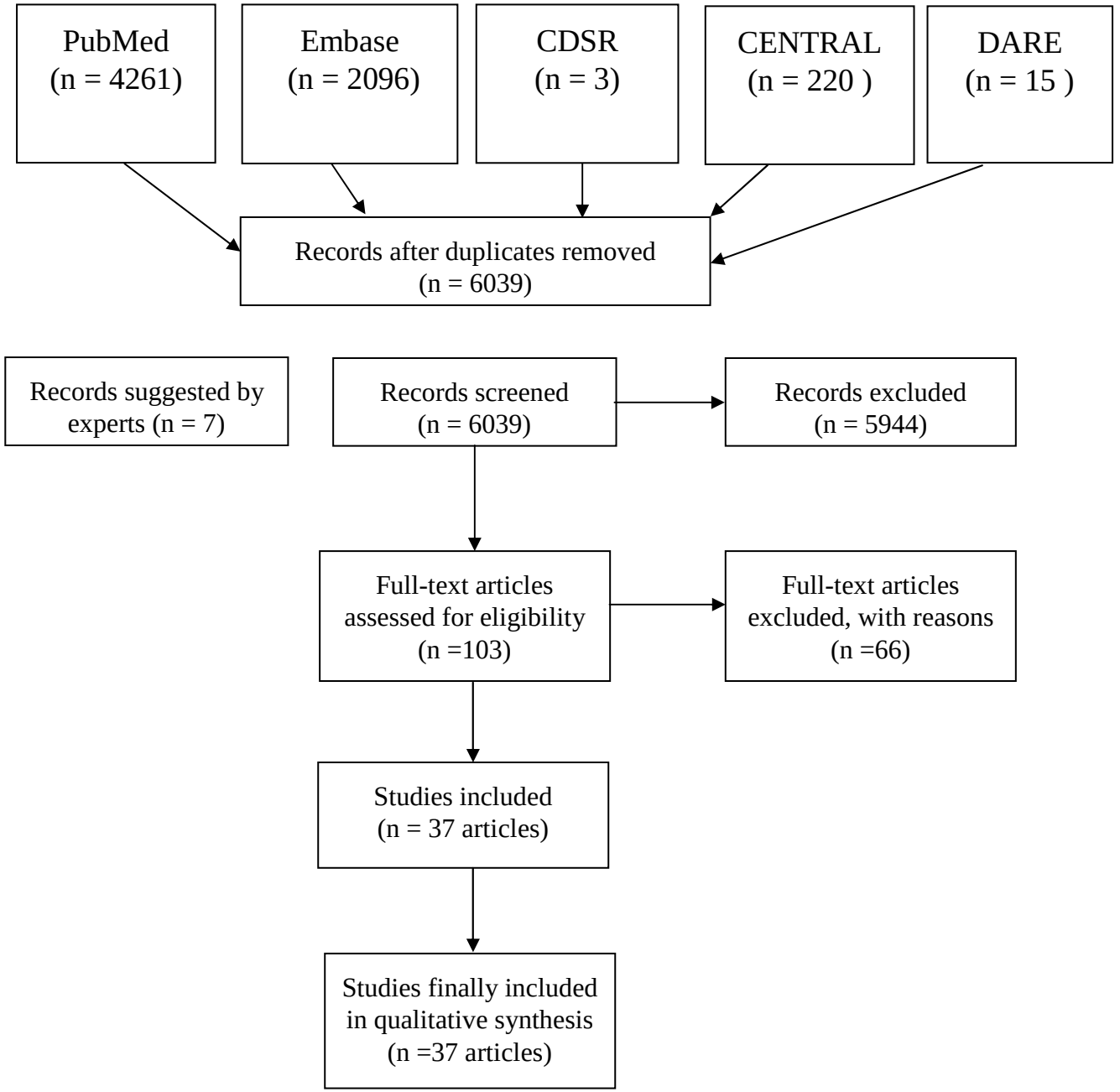
-e di studi di riproducibilità dei patologi nella diagnosi istologica (CB, EC).

Ricerca effettuata su PubMed, Embase e Cochrane library entro il 1/4/2014.

Outcome: 1)tasso di falsi positivi (TFP): percentuale di diagnosi istologiche riclassificate da maligne alla CB a benigne alla EC;

2)misclassificazione di diagnosi istologiche benigne in maligne (MBM), e statistica K negli studi di riproducibilità

# Materiali e Metodi



## *RISULTATI*

Sono stati inclusi complessivamente 36 studi.

### Outcome 1: CB vs EC

15 studi valutano TFP dell'esame istologico da CB, in casi diagnosticati nel 1990-2007.

In 10 studi che includono 42152 lesioni screen detected, TFP varia da 0% a 7%.

# Caratteristiche e risultati degli studi su CB vs EC-I

| Study | Any false positive at CB (false positive)/ all positives at CB | From invasive at CB to DCIS at surgery | From invasive at CB to benign at EC | From DCIS at CB to benign at EC |
|---|---|---|---|---|
| ***Screen detected lesions*** | | | | |
| Britton 1997 | * 2.08% (95% CI 0.25%- 7.32%) if malignant is defined as B4+B5<br>** 0% (97.5% CI 0%-4.02%) f malignant is defined as B5 only | 0% (97.5%CI 0%-3.93%) | *2.08% (95% CI 0.25%- 7.32%) (from B4 (probably malignant) to benign | |
| Dahlstrom 1996<br>Sutton 1997 | No false positive case 0% (97.5% CI 0%-6.98%)<br>1 patients chose not to have an excision | | | |
| Jackman 1994 | No false positive 0% (97.5% CI 0%-3.13%) | | | |
| Lifrange 1997 | 3 | | 1 (from mucinous carcinoma to fibroadenoma with myxoidstroma) | 2 (From DCIS to ADH) |
| Rakha 2009 | 0.02 % (95%CI 0.01%- 0.04%) | | 0.02%(%95%CI 0.01%- 0.04%) from B5b to B2 (fat necrosis) | 0.02% (95%CI 0.01%- 0.04%) (from B5a (DCIS) to either B3 or B4) |
| Seoudi 1998 | No false positive 0% (97.5% CI 0%-20.59%) | | | |
| Smyth 1994 | 7.14% (95% CI 0.18%-33.87%) | | | |
| Taft 1996 | 0.94% (95% CI 0.02%-5.14%) | | 0.94% (95% CI 0.02%-5.14%) from invasive to Radial scar with ADH | |
| Vega 1995 | No false positive case 0% (97.5% CI 0%-6.98%) | | | |
| Verkooijen 2002 | 1.88% ( 95% CI 0.86%; 3.53%) | 0.68% (95% CI 0.38%- 3.49%) | not enough details to allocate the further 5 FP to each group | not enough details to allocate the further 5 FP to each group |

0% campioni molto piccoli

*One patient refused operation, 3 patients treated with Tamoxifen alone
**2 patients treated with Tamoxifen alone, not surgery

# *RISULTATI-I*

## Outcome 2: MBM, statistica K

21 studi valutano la riproducibilità di 2 o più letture dello stesso vetrino.

Tra gli studi con *campioni consecutivi, casuali o stratificati di tutti i vetrini* MBM alla CB varia dallo 0.25% al 1.96% (3 studi), alla EC da 0.69% a 1.17% (2 studi). K varia alla CB da 0.83 a 0.98 (4 studi), e da 0.86 a 0.94 alla EC (3 studi).

# Caratteristiche e risultati degli studi su riproducibilità di due lutture dello stesso vetrino-Ia

| Study | 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) (%) | K value (when specified) or overall agreement or overall disagreement |
|---|---|---|---|---|---|
| **PATIENTS CONSECUTIVELY RECRUITED OR RANDOMLY SELECTED SAMPLES OR STRATIFIED SAMPLES** | | | | | |
| **Type of specimen: Core needle biopsies** | | | | | |
| ***Screen detected lesions*** | | | | | |
| Collins 2004 | 2.55% | 0.12% | 0.19% | 0.25% | K Overall 0.90 (95% CI 0.88%-0.92%)<br>Agreement: 96.06% (95% CI 95.11%-96.87%) |
| ***Screen and clinically detected lesions*** | | | | | |
| Verkooijen 2003 | 1.37% | 0.36% | 3.20% | 1.39% | Overall agreement 88.02%<br>(95% CI 85.42%- 90.31%)<br>K overall: 0.83 (95% CI 0.78–0.88) |
| Stang 2011 | NA | NA | NA | 1.96%<br>number of discordant diagnoses | K overall<br>five-level B-categorization scheme :<br>K : 0.89 (95% CI 0.86–0.91)<br>two-level B-categorization scheme<br>K 0.86 (95% CI: 0.83–0.90) |
| **Type of specimen: surgical specimens** | | | | | |
| ***Screen detected lesions*** | | | | | |
| Collins 2004 | 1.49% | 1.04% | 2.59% | 1.17% | Overall agreement: 92.62% (95% CI 90.22%–94.58%)<br>K Overall 0.89 (95% CI 0.86-0.92 ) |
| ***Screen and clinically detected lesions*** | | | | | |
| Verkooijen 2003 | 1.55% | 0% | 1.89% | 0.69% | Overall agreement 90.39% (95% CI 88.00%- 92.45%)<br>K overall:0.86 (95% CI 0.81–0.91) |

# *RISULTATI-II*

## Outcome 2: MBM, statistica K

Tra gli studi con *campioni arricchiti* MBM varia da 1.36% a 4.39% (5 studi), K da 0.57 a 0.86.

# Caratteristiche e risultati degli studi su riproducibilità di due lutture dello stesso vetrino-IIa

| Study | 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) (%) | K value (when specified) or overall agreement or overall disagreement |
|---|---|---|---|---|---|
| **ENRICHED SAMPLES** | | | | | |
| **Type of specimen: Core needle biopsies** | | | | | |
| ***Screen detected lesions*** | | | | | |
| Bianchi 2009 | NA | NA | NA | 4.39% (considering as benign B2 and B3 and malignant B4-B5) | Kappa overall =0.61 (range 0.31–0.88) |
| **Type of specimen: not reported** | | | | | |
| ***Screen detected lesions*** | | | | | |
| Bianchi 1994 | 3.47 % | 3.85% | 0% | 2.00% | K Overall =0.86 (range 0.65–1.0). |
| Sloane 1994 | 2.31% | 1.05% | 2.23% | 1.36% | K Overall<br>Coordinators 0.86<br>Non coordinators 0.78 |
| ***Not specified if clinically or screen detected lesions*** | | | | | |
| Beck 1985 | First circulation=0%<br>Second circulation=0% | First circulation= 0.89%<br>Second circulation= 0% | First circulation= 4.44%<br>Second circulation=4.17% | First circulation= 3.33%<br>Second circulation= 2.5% | K Overall =0.57 (value for both circulation and two borderline series combined) |
| Giardina 1998 | 0% | 4.38% | 3.19% | 3.68% | K Overall<br>Between pathologist=0.66 (range 0.57-0.76)<br>Between pathologist and the predominant diagnosis=0.786 (SE 0.27) |

# *RISULTATI-III*

## Outcome 2: MBM, statistica K

Tra gli studi con *casi selezionati per una seconda opinione* MBM varia da 0.29% a 12.16% (6 studi), K è riportato in 2 studi (0.48 e 0.5)

# Caratteristiche e risultati degli studi su riproducibilità di due lutture dello stesso vetrino-IIIa

| Study | 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) (%) | K value (when specified) or overall agreement or overall disagreement |
|---|---|---|---|---|---|
| **SECOND OPINION** | | | | | |
| **Type of specimen: core needle biopsies and surgical specimens** | | | | | |
| ***Not specified if clinically or screen detected lesions*** | | | | | |
| Marco 2013 | 5.59% | 11.76% | 0% | 0.98% | Overall agreement: 74.63%( 95% CI 68.00%- 80.44v) |
| Perez 2013 | 6.62% | 5.17% | 34.48% | 11.00% | Overall agreement: 83.25% ( 95% CI 77.49%-88.05%)<br>K overall: 0.5 |
| **Type of specimen: open surgery** | | | | | |
| ***Not specified if clinically or screen detected lesions*** | | | | | |
| Renshaw 2013 | 3/ not reported the number of malignant | NA | NA | NA | |
| **Type of specimen: not reported** | | | | | |
| ***Not specified if clinically or screen detected lesions*** | | | | | |
| De Almeida Salles 2008 | 12.00% | 7.75% | 23.26% | 12.16% | Overall agreement: 59.88% (95% CI 54.36%- 65.22%)<br>K overall: 0.48 |
| Newman 2006 | | | | 4.03% | Overall agreement=71.14%( 95% CI 63.16%- 78.26%) |
| Staradub 2002 | 2.03% | NA | NA | 0.29% | Overall agreement 80.35%( 95% CI 75.76%- 84.40%) |

# *CONCLUSIONI*

TFP e MBM possono contribuire in modo non trascurabile alla sovradiagnosi, variando TFP alla CB da 0% a 7% e MBM da 0.25% a 12.2%. Come atteso la concordanza tra patologi si riduce in studi con campioni arricchiti.

Prendendo in considerazione  la dimensione del campione e la qualità metologica, tra gli studi che valutano il TFP alla CB per lesioni screen detected Verkooijen 2002 può essere considerato uno studio informativo (TFP=1.88%).

L'impatto delle diagnosi istologiche false positive sulla sovradiagnosi non è trascurabile ed ha una rilevanza nella pratica clinica.

# *AGGIORNAMENTO DELLA RICERCA E MOTIVI DI ATTUALITÀ*

# *AGGIORNAMENTO della RICERCA*

Al 30/10/2015, Ulteriori 893 abstract, da cui sono stati selezionati per l'inclusione ed estrazione dei dati ulteriori: 1+7 (pubblicati tra Aprile 2014-Ottobre 2015). Tra cui

Original Investigation

## Diagnostic Concordance Among Pathologists Interpreting Breast Biopsy Specimens

Joann G. Elmore, MD, MPH; Gary M. Longton, MS; Patricia A. Carney, PhD; Berta M. Geller, EdD; Tracy Onega, PhD; Anna N. A. Tosteson, ScD; Heidi D. Nelson, MD, MPH; Margaret S. Pepe, PhD; Kimberly H. Allison, MD; Stuart J. Schnitt, MD; Frances P. O'Malley, MB; Donald L. Weaver, MD

*JAMA*. 2015;313(11):1122-1132. c

**OBJECTIVES** To quantify the magnitude of diagnostic disagreement among pathologists compared with a consensus panel reference diagnosis and to evaluate associated patient and pathologist characteristics.

## *Methods*

**240** breast biopsy specimens (**excisional or core needle**) randomly identified from a cohort of 19498 cases obtained from pathology registries in New Hampshire and Vermont. **Random, stratified sampling** was used to select cases based on the original pathologists' diagnoses.

Participants independently interpreted slides between **November 2011 and May 2014** from **4 test sets of 60** breast biopsies (1 slide per case), including 23 cases of invasive breast cancer, 73 ductal carcinoma in situ (DCIS), 72 with atypical hyperplasia (atypia), and 72 benign cases without atypia. Participants were **blinded** to the interpretations of other study pathologists and consensus panel members.

## *Results*

115 pathologists completed the study, providing 6900 individual case diagnoses.

Compared with the consensus-derived reference diagnosis,

the **overall concordance rate** of diagnostic interpretations of participating pathologists was **75.3**%(95% CI, 73.4%-77.0%; 5194 of 6900 interpretations)

| Consensus Reference Diagnosis | Pathologist Interpretation vs Consensus-Derived Reference Diagnosis, % (95% CI) | | | |
|---|---|---|---|---|
| | No. of Interpretations | Overall Concordance Rate | Overinterpretation Rate | Underinterpretation Rate |
| Benign without atypia | 2070 | 87 (85-89) | 13 (11-15) | |
| Atypia | 2070 | 48 (44-52) | 17 (15-21) | 35 (31-39) |
| DCIS | 2097 | 84 (82-86) | 3 (2-4) | 13 (12-15) |
| Invasive carcinoma | 663 | 96 (94-97) | | 4 (3-6) |

**Overinterpretation** of benign without atypia breast biopsies (13% among the 2070 interpretations for 72 benign without atypia cases in this study) may be occurring **more often than underinterpretation** of invasive breast cancer (4% among 663 interpretations for 23 cases in this study).

MBM=428/6900
=6.2%

Figure 3. Comparison of 115 Participating Pathologists' Interpretations vs the Consensus-Derived Reference Diagnosis for 6900 Total Case Interpretations[a]

| | | Participating Pathologists' Interpretation | | | | |
|---|---|---|---|---|---|---|
| | | Benign without atypia | Atypia | DCIS | Invasive carcinoma | Total |
| Consensus Reference Diagnosis[b] | Benign without atypia | 1803 | 200 | 46 | 21 | 2070 |
| | Atypia | 719 | 990 | 353 | 8 | 2070 |
| | DCIS | 133 | 146 | 1764 | 54 | 2097 |
| | Invasive carcinoma | 3 | 0 | 23 | 637 | 663 |
| | Total | 2658 | 1336 | 2186 | 720 | 6900 |

# Patient and Pathologist Characteristics Associated With Overinterpretation and Underinterpretation

Disagreement with the reference diagnosis was statistically significantly higher among biopsies from women with higher (n = 122) vs lower (n = 118) breast density on prior mammograms (overall concordance rate, 73% [95% CI, 71%-75%] for higher vs 77% [95% CI, 75%-80%] for lower, $P < .001$), and among pathologists who interpreted lower weekly case volumes ($P < .001$) or worked in smaller practices ($P = .034$) or nonacademic settings ($P = .007$).

The NEW ENGLAND JOURNAL of MEDICINE

# Reducing Diagnostic Errors — Why Now?

Dhruv Khullar, M.D., M.P.P., Ashish K. Jha, M.D., M.P.H., and Anupam B. Jena, M.D., Ph.D.

This article was published on September 23, 2015, at NEJM.org.

Although diagnosis has always been central to the practice of medicine and diagnostic errors have always been prevalent, **systematic efforts to measure these errors** and analyze their underpinnings **have been limited**, as compared with other quality and safety-improvement efforts.

….

But we would argue that **diagnostic errors are clinically and financially more costly today than ever before** and that they therefore require greater attention and more dedicated resources.

In the past, ….**More limited treatment options** for many conditions **meant less** likelihood of **iatrogenic harm from inappropriate interventions** and less potential for lost clinical benefit from appropriate ones.

# Reducing Diagnostic Errors — Why Now?

Dhruv Khullar, M.D., M.P.P., Ashish K. Jha, M.D., M.P.H., and Anupam B. Jena, M.D., Ph.D.

As **treatment options have become more effective and costly,** the clinical and financial costs of misdiagnosing a readily treatable condition are substantially greater.

….

Treating stage 4 colon cancer now costs more than three times what it costs to treat stage 1 disease.

As costly treatments for advanced disease become increasingly available, the costs of misdiagnosis — as well as those of overdiagnosis — can be expected to rise even further

…

**By failing to actively acknowledge and address the growing health and economic costs of diagnostic errors, we miss an important opportunity to provide better care for patients and realize better financial performance for health systems**.

# Is Breast Cancer Overdiagnosis Also Nested in Pathologic Misclassification?[1]

Catherine Colin, MD, PhD
Mojgan Devouassoux-Shisheboran, MD, PhD
Francesco Sardanelli, MD

***Radiology:*** Volume 273: Number 3—December 2014

mography (2). Is DCIS a cancer or not? Significant differences in perception were clearly shown by the answers given by 296 U.K. health-care professionals involved with the treatment of patients with DCIS (3). Breast cancer nurses, radiographers, physicians, radiologists, pathologists, and surgeons answered similarly. About 40% of these professional groups answered that DCIS is a cancer, while about 60% thought that DCIS is not a cancer. Only oncologists answered differently: More than 90% of them said that DCIS is not a cancer.

## The DCIS Nebula

Much of the discrepancies in classification arise from the criteria used for the distinction between atypical ductal hyperplasia (ADH) and DCIS (17). The threshold used to dif-

A second potential cause of misclassification and overdiagnosis is the presence or absence of microinvasion. DCIS

The most probable effect of this pathologic overestimation is, in the short term, a translation of a radiologic detection into a clinical overdiagnosis.

## Transition from DCIS into Invasive Cancer: Is the Linear Ductal Lesion Continuum Burning?

Two different models have been developed to explain the occurrence of invasive ductal cancers and the role of DCIS in the process. The two models

16q and gains of 1q). Although the "in situ low-grade breast neoplasia family" may constitute the precursor of low-grade breast cancers, its frequency of progression and role in the evolution of breast cancers remain unclear.

search. Research aimed at reducing DCIS overtreatment may be greatly facilitated by a new effort to rename and rethink these intraductal proliferative lesions in light of the most recent biologic and pathologic views.

## Not All Overdiagnoses Are Radiologic Diagnoses

Overdiagnosis and overtreatment raise complex and multifaceted issues. Difficulties in pathologic classifications and new biologic concepts lead to the possibility that an important causality of breast cancer overdiagnosis can be nested when analyzing and qualifying lesions. Difficulties and disagreements in pathologic interpretation and classification of intraductal proliferative lesions, as well as in diagnosing microinvasion, should be considered as potential sources of overdiagnosis. Thus, while detection (and overdetection) can be a radiologic issue, diagnosis (and overdiagnosis) should be shared by radiologists and pathologists. Sev-

# Grazie per l'attenzione!

# Caratteristiche e risultati degli studi su CB vs EC-II

| Study | Any false positive at CB (false positive)/ all positives at CB | From invasive at CB to DCIS at surgery | From invasive at CB to benign at EC | From DCIS at CB to benign at EC |
|---|---|---|---|---|
| ***Screen and clinically detected lesions*** | | | | |
| Frankel 2011 | No false positive 0%<br>(97.5% CI 0%-4.11%) | | 0%<br>(97.5% CI 0%-4.11%) | |
| Pijnappel 1997 | Palpable lesions: no false positive 0%<br>(97.5% CI 0%-15.44%)<br>Non palpable lesions: no false positive 0%<br>(97.5%CI 0%-8.22%) | | | |
| Wiratkapun 2010 | No false positive 0%<br>(97.5% CI 0%-24.71%) | | | |
| ***Not specified if clinically or screen detected lesions*** | | | | |
| Richter-Ehrenstein 2009 | 0.41%<br>(95%CI 0.05%- 1.47%) | 0.20% (95% CI 0.01%-1.14%)<br>from invasive to LIN | 0.20%<br>(95% CI 0.01%-1.14%)<br>from invasive to ADH | |
| Tse 2010 | No false positive 0%<br>(97.5% CI 0%-19.51%) | | | |