Mammography Report Peer Review: Critical Analysis of Two Years of Experience in a Private Hospital in São Paulo – Brazil
Authors

- Bruna Mayumi Takaki Tachibana
- Erica Elisangela Françolin Federicci
  - Larissa Muramoto Yano
  - Aline Lemgruber Prado Costa
  - Renata Feres
- Ana Carolina Dourado Mattos Raful
  - Monica Akahoshi Rudner
  - Ana Claudia Silveira Racy
- Marcelo Buarque de Gusmão Funari
  - Renato Leme de Moura Ribeiro

Disclosure: Nothing to disclose
Introduction

- Mammography is the exam of choice for breast cancer screening, since it can detect early stage breast cancers, reducing its mortality.

- Even with intense training and large experience, it is possible for a breast radiologist to miss lesions.

- To reduce the number of errors, a Quality Control (QC) Program must be implemented in an Imaging Department.

- Exam peer review with data analysis is one importante step of any comprehensive QC Program.

- We present the results of the peer review of mammography reports submitted to double reading during a period of 32 months.
Methods

- The data were collected from the peer review program from April 2014 to November 2016. All exams from this period were submitted to peer review evaluation during the reading process, in which the second reader analyzed agreement or disagreement. In some cases, a third reader was requested to reach a consensus.

- All exams were read by Breast Imaging radiologists.

- As part of the QC program we followed patients with disagreements for at least 6 months or correlated with histopathological data ensuring a feedback for the readers. We accomplished the follow up of 43% of cases with detection disagreements and 54% of cases with interpretation disagreements.

- The analysis focused mainly on significant disagreements related to detection and interpretation issues because they are directly related to radiologist’s performance.
Results

53,095 Reports

53,674 Reports

590 EXCLUDED*

49,826 (93.8%) NO DISAGREEMENTS

2,225 (4.2%) NON SIGNIFICANT DISAGREEMENTS

1,044 (2%) SIGNIFICANT DISAGREEMENTS

*Reports excluded due to missing data.
Results

Significant Disagreement Classification

- Grammar: 0.6%
- Lesion side error: 3.4%
- Exams’ technique: 4.8%
- Reporting semantics / Objectivity: 9.2%
- Detection: 19.0%
- No key images: 27.9%
- Interpretation: 35.2%
Detection

- The two most common lesions related to detection disagreements were calcifications (43%) and nodules (40%), followed by asymmetry (8%) and architectural distortion (7%).

- Among the followed-up cases, there were 45 reports in which the disagreement resulted in a BI-RADS category upgrade (from 1, 2 or 3 to 4 or 5) and the final result was:
  - Benign lesions in 56% (25);
  - High-risk lesions in 36% (16);
  - Malignant lesions in 9% (4).

- In other words: 45% of the cases with BI-RADS upgraded category were confirmed as high risk or malignant lesions.
Detection

- Studying the group with significant disagreement in Detection, we found a Positive Predictive Value (PPV) of 9% for BIRADS 4 or 5 lesions.

- This result is lower than the benchmark for general screening, according to ACR BI-RADS Atlas, 5th edition (between 20-50%).

- The lower PPV of this particular analyzed group is probably due to the size or aspect of the lesions, which were mostly subtle or small and not detected by the first reader.
FIRST READER: BENIGN FINDINGS;  
PEER REVIEW: ARCHITECTURAL DISTORTION;  
FINAL DIAGNOSIS: INVASIVE DUCTAL CARCINOMA.

FIRST READER: BENIGN CALCIFICATIONS.  
PEER REVIEW: CLUSTERED AMORPHOUS CALCIFICATIONS;  
FINAL DIAGNOSIS: DUCTAL CARCINOMA IN SITU.
Interpretation

- Interpretation related disagreements resulted in BI-RADS category upgrade in 48 cases (58%) and downgrade in 35 cases (42%);

- Among the BI-RADS category upgraded cases:
  - 56% were benign lesions (27);
  - 42% were high-risk lesions (20);
  - 2% were malignant lesions (1);

- Among the BI-RADS category downgraded cases:
  - 94% were benign lesions (33);
  - 3% were high-risk lesions (1);
  - 3% were malignant lesions (1).

DIAGNÓSTICO POR IMAGEM
Interpretation

- The most common causes of interpretation disagreements were:
  - Calcifications (40%);
  - Mass (25%);
  - Asymmetry (11%);
  - Follow up schedule (6%);
  - Final recommendation (5%).

- The two discordant malignant lesions were calcifications (one BI-RADS upgraded and the other downgraded by peer review).

- The data suggest that whenever calcifications or masses are involved, an extra care must be taken, since these are the most common causes of disagreements.
- FIRST READER: BENIGN CALCIFICATIONS;

- PEER REVIEW: A CLUSTER OF ROUND AND PUNCTATE CALCIFICATIONS, THAT INCREASED IN NUMBER COMPARED TO 2014 EXAM;

- FINAL DIAGNOSIS: DUCTAL CARCINOMA IN SITU.
Interpretation

• There were 13 cases with BI-RADS category upgrade related to masses.

• Nine of them resulted in benign lesions and four in high-risk lesions.

• On the other hand, there were 8 cases of masses with BI-RADS category downgrade and all of them were confirmed as benign lesions.

• The data analysis suggests that it may be necessary to review the benign and malignant lesion criteria particularly for mass type lesions, since there were no cases of malignant masses resulting from BI-RADS category upgrade.
Conclusions

- In modern times, quality control in medical care is an important issue.

- QC Programs must be implemented in any Imaging Department.

- Report peer review and data analysis are an important part of any QC Program. Our program detected and confirmed 41 additional malignant or high-risk lesions out of 53,095 reports (1 for each 1,295 reviewed) and downgraded only 2 out of the same basis.

- One big challenge nowadays is to convert all the information provided into learning opportunities to be shared among team members to prevent future mistakes.
REFERENCES


corresponding author: brunatakaki@gmail.com / bruna.takaki@einstein.br