# OVERVIEW

This document presents an overview of the most recent results of three proper[[1]](#footnote-1) ROC-curve fitting methods: the PROPROC1, CBM2 and the radiological search model3 (RSM) methods. **It corrects erroneous negative statements regarding PROPROC published in an earlier document[[2]](#footnote-2)**.

The three methods were applied to 14 datasets, indexed by **d** (**d** = 1, 2, ..., 14) With **I** modalities and **J** readers, each dataset yields **3IJ** ROC plots. Table 1 summarizes the datasets and acknowledges the sources. As an example, dataset 7 with 5 modalities and 7 readers yields 3 x 35 plots. The total number of individual modality-reader combinations is 236, i.e., , meaning there are 3 x 236 plots in all.

All plots are in document RSM PROPROC CBM 236 Plots.pdf. The filled circles are the operating points. The black line is the RSM fit. It always includes a dotted line extension from uppermost non-trivial point to (1,1). The blue line is the CBM fit. It extends continuously to (1,1). The red line is the PROPROC fit. It too extends continuously to (1,1). Each ROC plot is labeled by the dataset number, the modality-index i and the reader-index j. For example, D1, i = 2, j = 2 refers to dataset 1, modality 2 and reader 2.

Table 1: This table lists the summary characteristics of the datasets used in this book. The dataset type is ROC, FROC or LROC. [I = # modalities, J = # readers, K1 = number of non-diseased cases, K2 = number of diseased cases; K = total # cases.]

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Dataset # d** | **Dataset Name** | **Data**  **Type** | **I** | **J** | **K1** | **K2** | **K** | **Description** |
| 1 | TONY4 | FROC | 2 | 5 | 96 | 89 | 185 | Digital breast tomosynthesis vs. mammography |
| 2 | VD5 | ROC | 2 | 5 | 69 | 45 | 114 | Cine vs. SE MRI for aortic dissection |
| 3 | FR6 | ROC | 2 | 4 | 33 | 67 | 100 | Digital vs. analog pediatric chest |
| 4 | FED7 | FROC | 5 | 4 | 100 | 100 | 200 | Image processing in mammography: FROC |
| 5 | JT8 | FROC | 2 | 9 | 45 | 47 | 92 | Nodule detection in an thorax CT phantom |
| 6 | MAG9 | FROC | 2 | 4 | 47 | 42 | 89 | Tomosynthesis Vs. Radiography Pulmonary Nodules |
| 7 | OPT10 | FROC | 5 | 7 | 81 | 81 | 162 | Calcification detection in digital mammography |
| 8 | PEN11 | FROC | 5 | 5 | 48 | 64 | 112 | Image compression in mammography |
| 9 | NICO12 | LROC | 1 | 10 | 120 | 80 | 200 | Standalone CAD (modality 1) vs. 9 radiologists |
| 10 | RUS13 | FROC | 3 | 8 | 50 | 40 | 90 | Lesion detection in digital mammography |
| 11 | DOB114 | FROC | 4 | 5 | 43 | 115 | 158 | Tomosynthesis, Dual-Energy & Conventional Chest |
| 12 | DOB214 | ROC | 4 | 5 | 64 | 88 | 152 | do: |
| 13 | DOB314 | FROC | 4 | 5 | 52 | 106 | 158 | do: |
| 14 | FZR15 | ROC | 2 | 4 | 100 | 100 | 200 | Image processing in mammography: ROC |

**None of this will make much sense unless one takes the time to understand three basic concepts**.

1. The **area under the ROC** curve measures the ability of the observer to **discriminate diseased from non-diseased cases**. Perfect discrimination yields ROC-AUC = 1 while guessing performance yields ROC-AUC = 0.5. ROC-AUC gives no information on what is limiting performance: search performance or lesion-classification performance.
2. **Search performance** is the ability of the observer to find true lesions while avoiding false lesions. Whether a found lesion is marked depends on the reporting threshold. It is measured by a quantity denoted **S**.
3. Having found a suspicious lesion, **lesion classification** performance is the ability to **correctly classify it is malignant or benign**. It is denoted by , ranging from 0.5 (chance level ability) to 1 (perfect).
4. None of the conventional ROC models measures search performance. Lesion-classification performance is measured for some CAD systems (termed CADx).

The primary conclusion of this document is that search performance is the "bottleneck", averaging about 17%, while lesion-classification performance is about 89%. Unless researchers recognize this fact, and continue to measure ROC performance (as suggested by the FDA and some statisticians) their efforts will not be productive. A fundamental rule of science is to first focus on the weak-link.

# Definitions

* **S** = search performance; denoted in book, ability to find lesions while avoiding non-lesions: 0 < S < 1.
* **C**= lesion classification performance; denoted  in book, ability to discriminate between lesions and non-lesions: 0.5 <  < 1.
* **A** = ROC-AUC, denoted  in book, ability to discriminate between non-diseased and diseased cases.
* **S**, **C** and **A** are related to 3 RSM-parameters  and the observed lesion distribution vector , which were determined by MLE for each ROC dataset. The value  is not a parameter, but the observed fraction of lesions/case vector, i.e., the fraction of cases with one lesion, the fraction with two lesions, etc., which sums to unity.

# Overall conclusions

* Search performance is low (0.17) compared to lesion-classification performance (0.89). **Search performance is by far the weak link in observer performance**. It is not measurable under any conventional ROC model, i.e., those that generate decision variable samples on every image. It is measurable using RSM fitting model.
* There exists a **strong inverse correlation** between **C** and **S** and weaker positive correlations between **A** vs. **S** and **A** vs. **C**. Observers tend to compensate for deficiencies in search by greater performance in lesion-classification, and vice-versa. The ideal observer represents an upper limit on net ROC performance,

Overall conclusions from Fig. 02:

* All three methods yield almost identical AUCs. PROPROC AUC is about 1% larger than RSM AUC, while CBM AUC is about 1% smaller than RSM AUC.
* Consistent with each proper-ROC method as a realization of an ideal observer.
* RSM-mu and CBM-mu parameters are correlated, consistent with their physical meanings.
* RSM-nu and CBM-alpha parameters are correlated, consistent with their physical meanings.
* Therefore, the CBM and RSM models are mutually reinforcing.
* For degenerate datasets PROPROC yields gross overestimates of performance 16. This problem is readily fixed but unfortunately, this important work by Metz and colleagues is no longer being supported.

# Dataset D01: TONY Data

## Intra-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |

Under **Intra-correlations**, are three plots: one shows the relation between lesion-classification performance **C** and search performance **S**, the second showing that between **A** vs. **S** and finally, that between **A** vs. **C**.

## Inter-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **PRO vs. RSM AUCs slope = 1.018522 , r2 = 0.9990286** | **CBM vs. RSM AUCs slope = 1.013233 , r2 = 0.9994436** |  |

Under **Inter-correlations**, are two plots: one shows the relation PROPROC-AUC and RSM-AUC and the other shows the relation PROPROC-AUC and RSM-AUC. Listed below each plot are the slope of a linear fit constrained to go through the origin and the R2 of the fit.

# D02: VD Data

## Intra-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |

## Inter-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **PRO vs. RSM AUCs slope = 1.000744 , r2 = 0.9996271** | **CBM vs. RSM AUCs slope = 0.9965077 , r2 = 0.999637** |  |

# D03: FR Data

## Intra-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |

dsfakjlaskl .

## Inter-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **PRO vs. RSM AUCs slope = 1.000198 , r2 = 0.9999957** | **CBM vs. RSM AUCs slope = 0.9983922 , r2 = 0.9999967** |  |

# D04: FED data

## Intra-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |

## Inter-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **PRO vs. RSM AUCs slope = 1.020122 , r2 = 0.9997407** | **CBM vs. RSM AUCs slope = 1.006134 , r2 = 0.9998828** |  |

# D05: JT data

## Intra-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |

## Inter-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **PRO vs. RSM AUCs slope = 1.016888 , r2 = 0.9994528** | **CBM vs. RSM AUCs slope = 1.003703 , r2 = 0.9999301** |  |

# D06: MAG data

## Intra-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |

## Inter-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **PRO vs. RSM AUCs slope = 1.018522 , r2 = 0.9990286** | **CBM vs. RSM AUCs slope = 1.013233 , r2 = 0.9994436** |  |

# D07: OPT data

## Intra-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |

## Inter-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **PRO vs. RSM AUCs slope = 1.024632 , r2 = 0.9994263** | **CBM vs. RSM AUCs slope = 1.013228 , r2 = 0.9996222** |  |

# D08: PEN data

## Intra-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |

## Inter-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **PRO vs. RSM AUCs slope = 1.011926 , r2 = 0.9995145** | **CBM vs. RSM AUCs slope = 1.00208 , r2 = 0.9999539** |  |

# D09: NICO data

## Intra-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |

## Inter-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **PRO vs. RSM AUCs slope = 1.008892 , r2 = 0.9998246** | **CBM vs. RSM AUCs slope = 1.006252 , r2 = 0.999812** |  |

# D10: RUS data

## Intra-correlations

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | |  |
|  | |  | |

## Inter-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **PRO vs. RSM AUCs slope = 1.009527 , r2 = 0.9992249** | **CBM vs. RSM AUCs slope = 1.003992 , r2 = 0.999374** |  |

# D11: DOB1 data

## Intra-correlations

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | |  |
|  | |  | |

## Inter-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **PRO vs. RSM AUCs slope = 1.081471 , r2 = 0.9968004** | **CBM vs. RSM AUCs slope = 1.038696 , r2 = 0.9983992** |  |

# D12: DOB2 data

## Intra-correlations

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | |  |
|  | |  | |

## Inter-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **PRO vs. RSM AUCs slope = 1.063601 , r2 = 0.998276** | **CBM vs. RSM AUCs slope = 1.028398 , r2 = 0.9989182** |  |

# D13: DOB3 data

## Intra-correlations

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | |  |
|  | |  | |

## Inter-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **PRO vs. RSM AUCs slope = 1.061524 , r2 = 0.9971542** | **CBM vs. RSM AUCs slope = 1.026581 , r2 = 0.9990437** |  |

# D14: FZREAL data

## Intra-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |

## Inter-correlations

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **PRO vs. RSM AUCs slope = 1.002708 , r2 = 0.9999087** | **CBM vs. RSM AUCs slope = 0.9996302 , r2 = 0.9999998** |  |

# SUMMARY 1: Bootstrap analysis intra-correlations (RSM model-based correlations)

|  |  |  |
| --- | --- | --- |
|  |  |  |
| Avg. rhoSC = -0.564 (-0.658, -0.387) | Avg. rhoSA = 0.295 (0.119, 0.414) | Avg rhoAC = 0.013 (-0.109, 0.262) |
|  |  |  |
| Avg S = 0.166 (0.14, 0.19) | Avg C = 0.895 (0.88, 0.91) | Avg A = 0.803 (0.79, 0.81) |

Fig. 01: Listed under "Bootstrap analysis intra-correlations (RSM model-based correlations)" are three bootstrap histograms of **S**, **C** and **A**, the average and 95% confidence intervals for each quantity. These are followed by histograms of the Pearson correlations between **S** and **C**, between **S** and **A** andbetween **A** and **C**, the corresponding averages and empirical 95% confidence intervals.

> source('~/book3/04 C FROC paradigm/C18 RSM Fitting/software/mainIntraCorrelationsCI.R')

Avg S = 0.1660226 , Avg C = 0.8953426 , Avg A = 0.8034145 Avg rhoSC = -0.5642952 , Avg rhoSA = 0.2948044 , Avg rhoAC = 0.01320294

The empirical 95% CI of S is 0.139365119683602, 0.19207779193

The empirical 95% CI of C is 0.878824685458566, 0.911426380681954

The empirical 95% CI of A is 0.79463405161037, 0.812838504948535

The empirical 95% CI of the correlation between S and C is -0.658155337918353, -0.387393202749413

The empirical 95% CI of the correlation between S and A is 0.119454039265007, 0.414429009317891

The empirical 95% CI of the correlation between A and C is -0.109161489323403, 0.262373375623688

# SUMMARY 2: Bootstrap analysis inter-correlations (slope plot summary)

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |
| Avg. slopeProRsm = 1.010094 (1.0070, 1.0139) | Avg. slopeCbmRsm = 0.9938181 (0.99161, 0.99612) |  |

Fig. 02: Listed under "Bootstrap analysis inter-correlations (slope plot summary)" are bootstrap histograms of the slopes of constrained linear fits between PROPROC and RSM AUCs, and between CBM and RSM AUCs, the averages and 95% confidence intervals.

> source('~/book3/04 C FROC paradigm/C18 RSM Fitting/software/mainInterCorrelationsCI.R')

rhoMuRsmMuCbm[f]= 0.7529951 , rhoNupRsmAlphaCbm[f]= 0.9724744

rhoMuRsmMuCbm[f]= 0.9640124 , rhoNupRsmAlphaCbm[f]= 0.9860099

rhoMuRsmMuCbm[f]= 0.09006163 , rhoNupRsmAlphaCbm[f]= 0.9157216

rhoMuRsmMuCbm[f]= 0.8876381 , rhoNupRsmAlphaCbm[f]= 0.9683928

rhoMuRsmMuCbm[f]= 0.925323 , rhoNupRsmAlphaCbm[f]= 0.9530313

rhoMuRsmMuCbm[f]= 0.9906749 , rhoNupRsmAlphaCbm[f]= 0.9783978

rhoMuRsmMuCbm[f]= 0.8028458 , rhoNupRsmAlphaCbm[f]= 0.8257765

rhoMuRsmMuCbm[f]= 0.9402888 , rhoNupRsmAlphaCbm[f]= 0.9565226

rhoMuRsmMuCbm[f]= 0.6783968 , rhoNupRsmAlphaCbm[f]= 0.9786479

rhoMuRsmMuCbm[f]= 0.8157921 , rhoNupRsmAlphaCbm[f]= 0.9644153

rhoMuRsmMuCbm[f]= 0.9527844 , rhoNupRsmAlphaCbm[f]= 0.94351

rhoMuRsmMuCbm[f]= 0.9677316 , rhoNupRsmAlphaCbm[f]= 0.9739261

rhoMuRsmMuCbm[f]= 0.9743165 , rhoNupRsmAlphaCbm[f]= 0.766374

rhoMuRsmMuCbm[f]= 0.9892122 , rhoNupRsmAlphaCbm[f]= 0.9908596

avg aucRsm = 0.8034145

, avg aucPro = 0.8109336

, avg aucCbm = 0.7992618

, slopeCbmRsm = 0.9938181

, avg R2CbmRsm = 0.999622

, slopeProRsm = 1.009171

, avg R2ProRsm = 0.9994403

, avg rhoMuRsmMuCbm = 0.8380052

, avg rhoNupRsmAlphaCbm = 0.9410043

The empirical 95% CI of avgAucRsm is 0.79463405161037, 0.812838504948535

The empirical 95% CI of avgAucPro is 0.802527813248711, 0.819638747111847

The empirical 95% CI of avgAucCbm is 0.790884163650715, 0.807888126726429

The empirical 95% CI of slopeProRsm is 1.00627592872924, 1.01255788148232

The empirical 95% CI of slopeCbmRsm is 0.991605598776276, 0.996122546285231

The empirical 95% CI of rhoMuRsmMuCbm is 0.717689195882282, 0.897340474301577

The empirical 95% CI of rhoNupRsmAlphaCbm is 0.899145670914886, 0.952872459031796

# SUMMARY 3: Bootstrap analysis inter-correlations (correlations between RSM and CBM parameters)

|  |  |  |
| --- | --- | --- |
|  |  |  |
| Avg. rhoMuRsmMuCbm = 0.83801 (0.7177, 0.8973) | Avg. rhoNupRsmAlphaCbm = 0.94100 (0.8991, 0.9529) |  |

Fig. 03: Listed under "Bootstrap analysis inter-correlations (correlations between RSM and CBM parameters)" are two plots of the bootstrap histograms of the Pearson correlation between RSM-mu and CBM mu-parameters and between RSM-nu and CBM alpha-parameters, the averages and 95% confidence intervals.

> source('~/book3/04 C FROC paradigm/C18 RSM Fitting/software/mainInterCorrelationScatterPlots.R')

TONY PRO vs. RSM AUCs slope = 0.9941862 , r2 = 0.9999593

TONY CBM vs. RSM AUCs slope = 0.989041 , r2 = 0.9998697

VD PRO vs. RSM AUCs slope = 1.008517 , r2 = 0.9997585

VD CBM vs. RSM AUCs slope = 1.001086 , r2 = 0.9999805

FR PRO vs. RSM AUCs slope = 0.9993238 , r2 = 0.999998

FR CBM vs. RSM AUCs slope = 0.9975152 , r2 = 0.9999925

FED PRO vs. RSM AUCs slope = 1.013062 , r2 = 0.9997307

FED CBM vs. RSM AUCs slope = 0.9979407 , r2 = 0.9999461

JT PRO vs. RSM AUCs slope = 1.007523 , r2 = 0.9995883

JT CBM vs. RSM AUCs slope = 0.9961413 , r2 = 0.9999678

MAG PRO vs. RSM AUCs slope = 1.03719 , r2 = 0.9983294

MAG CBM vs. RSM AUCs slope = 0.9953975 , r2 = 0.9999493

OPT PRO vs. RSM AUCs slope = 1.014075 , r2 = 0.9995914

OPT CBM vs. RSM AUCs slope = 1.002708 , r2 = 0.9997552

PEN PRO vs. RSM AUCs slope = 1.010077 , r2 = 0.9993828

PEN CBM vs. RSM AUCs slope = 0.9978563 , r2 = 0.9999665

NICO PRO vs. RSM AUCs slope = 1.003058 , r2 = 0.9999423

NICO CBM vs. RSM AUCs slope = 1.001027 , r2 = 0.9999722

RUS PRO vs. RSM AUCs slope = 0.9985236 , r2 = 0.9999161

RUS CBM vs. RSM AUCs slope = 0.9931587 , r2 = 0.9998827

DOB1 PRO vs. RSM AUCs slope = 1.019294 , r2 = 0.9991465

DOB1 CBM vs. RSM AUCs slope = 0.9835848 , r2 = 0.9992774

DOB2 PRO vs. RSM AUCs slope = 1.006628 , r2 = 0.9985419

DOB2 CBM vs. RSM AUCs slope = 0.9728445 , r2 = 0.9976341

DOB3 PRO vs. RSM AUCs slope = 1.014241 , r2 = 0.9983697

DOB3 CBM vs. RSM AUCs slope = 0.9855389 , r2 = 0.9985146

FZR PRO vs. RSM AUCs slope = 1.002695 , r2 = 0.9999092

FZR CBM vs. RSM AUCs slope = 0.9996141 , r2 = 0.9999999

> source('~/book3/04 C FROC paradigm/C18 RSM Fitting/software/mainIntraCorrelationScatterPlots.R')

Avg S[f] = 0.2705408 , Avg C[f] = 0.8371479 , Avg A[f] = 0.8176612 , RhoSC[f] = -0.2894877 , RhoSA[f] = 0.6634098 , RhoAC[f] = 0.2612029

Avg S[f] = 0.440869 , Avg C[f] = 0.9724411 , Avg A[f] = 0.9174 , RhoSC[f] = -0.4821393 , RhoSA[f] = 0.3279341 , RhoAC[f] = 0.2370426

Avg S[f] = 0.08705611 , Avg C[f] = 0.9676417 , Avg A[f] = 0.861533 , RhoSC[f] = -0.7881024 , RhoSA[f] = 0.570703 , RhoAC[f] = -0.3749476

Avg S[f] = 0.298025 , Avg C[f] = 0.9464675 , Avg A[f] = 0.8484512 , RhoSC[f] = -0.5884049 , RhoSA[f] = -0.0926666 , RhoAC[f] = -0.347433

Avg S[f] = 0.1864557 , Avg C[f] = 0.9351482 , Avg A[f] = 0.8824647 , RhoSC[f] = -0.5428038 , RhoSA[f] = 0.02324078 , RhoAC[f] = 0.2708509

Avg S[f] = 0.1267499 , Avg C[f] = 0.9431972 , Avg A[f] = 0.7019879 , RhoSC[f] = -0.213541 , RhoSA[f] = -0.3780073 , RhoAC[f] = -0.03204039

Avg S[f] = 0.3635329 , Avg C[f] = 0.9271014 , Avg A[f] = 0.8416527 , RhoSC[f] = -0.2262033 , RhoSA[f] = 0.253856 , RhoAC[f] = 0.4275496

Avg S[f] = 0.2005551 , Avg C[f] = 0.9532861 , Avg A[f] = 0.8256936 , RhoSC[f] = -0.2434242 , RhoSA[f] = 0.1571625 , RhoAC[f] = 0.4155096

Avg S[f] = 0.2124423 , Avg C[f] = 0.9262345 , Avg A[f] = 0.8547918 , RhoSC[f] = -0.6669007 , RhoSA[f] = -0.4271226 , RhoAC[f] = 0.3669759

Avg S[f] = 0.1954653 , Avg C[f] = 0.8255456 , Avg A[f] = 0.7843663 , RhoSC[f] = -0.5746013 , RhoSA[f] = 0.2564148 , RhoAC[f] = 0.2884126

Avg S[f] = 0.2927199 , Avg C[f] = 0.7824251 , Avg A[f] = 0.7220886 , RhoSC[f] = -0.4728074 , RhoSA[f] = 0.8236548 , RhoAC[f] = -0.1988058

Avg S[f] = 0.2585526 , Avg C[f] = 0.7558191 , Avg A[f] = 0.693241 , RhoSC[f] = -0.5231941 , RhoSA[f] = 0.7071523 , RhoAC[f] = -0.3035116

Avg S[f] = 0.1821415 , Avg C[f] = 0.7291555 , Avg A[f] = 0.6249957 , RhoSC[f] = -0.3728071 , RhoSA[f] = 0.9181937 , RhoAC[f] = -0.3321823

Avg S[f] = 0.1558031 , Avg C[f] = 0.9690183 , Avg A[f] = 0.8815373 , RhoSC[f] = -0.3756476 , RhoSA[f] = -0.4638648 , RhoAC[f] = 0.6025151

# References

1. Metz CE, Pan X. Proper Binormal ROC Curves: Theory and Maximum-Likelihood Estimation. *J Math Psychol.* 1999;43(1):1-33.

2. Dorfman DD, Berbaum KS. A contaminated binormal model for ROC data: Part II. A formal model. *Acad Radiol.* 2000;7(6):427-437.

3. Chakraborty DP. *OBSERVER PERFORMANCE METHODS FOR DIAGNOSTIC IMAGING - Foundations, Modeling, and Applications with R-Based Examples.* Taylor-Francis LLC; 2017 (under production, expected to be available by Dec 15, 2017).

4. Svahn T, Andersson I, Chakraborty D, et al. The Diagnostic Accuracy of Dual-View Digital Mammography, Single-View Breast Tomosynthesis and a Dual-View Combination of Breast Tomosynthesis and Digital Mammography in a Free-response Observer Performance Study. *Radiat Prot Dosimetry.* 2010;139:113–117.

5. Van Dyke CW, White RD, Obuchowski NA, Geisinger MA, Lorig RJ, Meziane MA. Cine MRI in the diagnosis of thoracic aortic dissection. *79th RSNA Meetings.* 1993.

6. Franken EA, Jr., Berbaum KS, Marley SM, et al. Evaluation of a Digital Workstation for Interpreting Neonatal Examinations: A Receiver Operating Characteristic Study. *Investigative Radiology.* 1992;27(9):732-737.

7. Zanca F, Jacobs J, Van Ongeval C, et al. Evaluation of clinical image processing algorithms used in digital mammography. *Medical Physics.* 2009;36(3):765-775.

8. Thompson JD, Chakraborty DP, Szczepura K, et al. Effect of reconstruction methods and x-ray tube current-time product on nodule detection in an anthropomorphic thorax phantom: a crossed-modality JAFROC observer study. *Medical Physics.* 2016;43(3):1265-1274.

9. Vikgren J, Zachrisson S, Svalkvist A, et al. Comparison of Chest Tomosynthesis and Chest Radiography for Detection of Pulmonary Nodules: Human Observer Study of Clinical Cases. *Radiology.* 2008;249(3):1034-1041.

10. Warren LM, Mackenzie A, Cooke J, et al. Effect of image quality on calcification detection in digital mammography. *Medical Physics.* 2012;39(6):3202-3213.

11. Penedo M, Souto M, Tahoces PG, et al. Free-Response Receiver Operating Characteristic Evaluation of Lossy JPEG2000 and Object-based Set Partitioning in Hierarchical Trees Compression of Digitized Mammograms. *Radiology.* 2005;237(2):450-457.

12. Hupse R, Samulski M, Lobbes M, et al. Standalone computer-aided detection compared to radiologists’ performance for the detection of mammographic masses. *Eur Radiol.* 2013;23(1):93-100.

13. Ruschin M, Timberg P, Bath M, et al. Dose dependence of mass and microcalcification detection in digital mammography: free response human observer studies. *Med Phys.* 2007;34:400 - 407.

14. Dobbins III JT, McAdams HP, Sabol JM, et al. Multi-Institutional Evaluation of Digital Tomosynthesis, Dual-Energy Radiography, and Conventional Chest Radiography for the Detection and Management of Pulmonary Nodules. *Radiology.* 2016;282(1):236-250.

15. Zanca F, Hillis SL, Claus F, et al. Correlation of free-response and receiver-operating-characteristic area-under-the-curve estimates: Results from independently conducted FROC/ROC studies in mammography. *Med Phys.* 2012;39(10):5917-5929.

16. Zhai X, Chakraborty DP. A bivariate contaminated binormal model for robust fitting of proper ROC curves to a pair of correlated, possibly degenerate, ROC datasets. *Med Phys.* 2017;44(3):in press.

1. As the operating point moves up the curve, a proper ROC curve has monotonically decreasing slope; it will not cross the chance diagonal nor will it show a "hook" near the top-right corner, as usually observed with the binormal model. [↑](#footnote-ref-1)
2. <https://www.researchgate.net/publication/317087463_Quantifying_search_lesion_classification_and_case_classification_performances>. [↑](#footnote-ref-2)