SSJ PUBLICATIONS

Journal of Integrated Human Health Sciences

ISSN-3143-8814

Research Article

# Diagnostic Information Is Measurement-Domain Dependent Reconstruction Entropy and Cross-Domain Discordance in Breast Fine-Needle Aspiration Cytology A proof-of-concept Operational Measurement-Domain Analysis Using the Wisconsin Diagnostic Breast Cancer Dataset

**Ahed J Alkhatib**[1*]

[1]*Department of Legal Medicine, Toxicology and Forensic Medicine, Jordan University of Science & Technology, Jordan*

***Correspondence:** Alkhatib A.J, Department of Legal Medicine, Toxicology and Forensic Medicine, Jordan University of Science & Technology, Jordan; E-mail: ajalkhatib@just.edu.jo

**Received date:** 14 Apr 2026: **Accepted date:** 05 May 2026: **Published date:** 12 May 2026

**Copyright:** © 2026 Alkhatib A.J. This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits use, sharing, adaptation, distribution, and reproduction in any medium or format, provided the original author(s) and source are properly credited.

## Abstract

**Background:** Digital cytology turns intricate cellular structure into quantifiable characteristics, but when it comes down to it, various summaries from a specific biological sample do not all impart the same level of diagnostic information.

**Objective:** To examine if it is possible to diagnose breast cancer based on its accessible nuclear-feature summary. Moreover, to quantify any cross-domain discordance along with uncertainty reduction.

**Methods:** The Wisconsin Diagnostic Breast Cancer data set contains 569 fine needle aspirate cases with 357 benign and 212 malignant cases having 30 image-derived nuclear features. The mean, SE (standard error) and "worst" values, where the "worst" values are defined in the source dataset by the mean of three largest values, are represented by ten morphological characteristics. Standardized L2-regularized logistic regression with repeated stratified 10-fold cross-validation (20 repeats) evaluated seven measurement domain models. Twenty out-of-fold probabilities for each case, being averaged before evaluation. The evaluation was done based on discrimination, accuracy, sensitivity, specificity, Brier score, binary prediction entropy, domain discordance, exact paired comparisons, and incremental entropy reduction.

**Results:** The Least informative domain was SE (AUC 0.943; accuracy 87.7 %), while the Worst domain was AUC 0.992; accuracy 97.7 %. AUC score of 0.995 and an accuracy of 97.9%. The average prediction entropy decreased from 0.455 bits for SE to 0.162 for Worst and 0.125 for all features. In 85 of the 569 total cases (14.9%), the Mean, SE, and Worst classifications disagreed. The accuracies of discordant cases were 63.5%, 22.4%, 89.4% and 91.8% for Mean, SE, Worst and after integration respectively.



**Citation:** Alkhatib A.J (2026). "Diagnostic Information Is Measurement-Domain Dependent Reconstruction Entropy and Cross-Domain Discordance in Breast Fine-Needle Aspiration Cytology A proof-of-concept Operational Measurement-Domain Analysis Using the Wisconsin Diagnostic Breast Cancer Dataset" *Journal of Integrated Human Health Sciences, 2*(1):1-13.

SSJ PUBLICATIONS
Journal of Integrated Human Health Sciences
ISSN-3143-8814
Research Article

**Conclusion:** The manner in which nuclear morphology is summed up determines the diagnostic information present in classic cytology datasets, which are highly dependent on the answer of this question. Their findings suggest an operational distinction between the biological specimen and the measurement domain through which it is represented and highlight that the model entropy is not an uncertainty measure that is biological but model dependent.

**Keywords:** Breast cancer, Fine-needle aspiration cytology, Digital cytology, Diagnostic uncertainty, Prediction entropy, Operational measurement domain, Machine learning, Wisconsin Diagnostic Breast Cancer

## Introduction

Fine needle aspiration cytology (FNAC) has long been a quick and easeful way to evaluate breast masses. The benefit of diagnosis will depend on adequate sampling, recognizable cytomorphology, and appropriate interpretation. Furthermore, contemporary reporting systems continue to emphasize diagnostic performance and uncertainty [1,2]. Even though, nowadays due to technological advancement core needle biopsy is the most performed procedure worldwide, FNAC is still clinically relevant procedure if performed and interpreted according to a diagnostic pathway. The Wisconsin Diagnostic Breast Cancer (WDBC) dataset is historically important for it translated breast FNA nuclear morphology into computer-generated measurements. Furthermore, it demonstrated that quantitative image features could differentiate benign from malignant lesions with a high degree of accuracy [3-6]. The original set of data has 569 cases and ten nuclear characteristics that describe size, texture, contour, shape irregularity and related morphology. It is critical in my current study that each characteristic has three different summaries for the purpose of analysis: the mean of values, the standard error of the mean and a “worst” summary which is based on the mean of the three highest values.

Because of this architecture, we have an extraordinary opportunity to ask a question that is different from merely maximizing classification accuracy. The biological specimen is identical, but the available numerical representation varies depending on the summary operator used. The mean, SE and Worst are therefore three measurement-summary domains of the same cytology source. The unequal diagnostic value of propositions means that diagnostic reconstruction depends not only on the type of lesion but also on which representation is available.

More and more relevant with the advent of computational pathology and clinical AI. Although contemporary systems may include morphological features, imaging features, molecular information, or features from multiple views, high discrimination alone does not guarantee safe clinical prediction. The issue of calibration, prediction uncertainty, applicability and transparent validation has gained increasing prominence in methodological guidance [7-10]. The use of an AI which is not effective and legalised in cytopathology investigations is growing quickly. And although other areas of technological development grow, analytical validity and clinical translation remain core issues [11].

As proof of concept, WDBC is utilized in this study in a measurement-domain framework. It does not offer a novel biological mechanism and does not consider the summaries of the three features to be independent tissue microregions. Rather, it investigates whether variations in the interpretation of the same cytological specimen lead to different diagnostic outcomes.



**Citation:** Alkhatib A.J (2026). “Diagnostic Information Is Measurement-Domain Dependent Reconstruction Entropy and Cross-Domain Discordance in Breast Fine-Needle Aspiration Cytology A proof-of-concept Operational Measurement-Domain Analysis Using the Wisconsin Diagnostic Breast Cancer Dataset” *Journal of Integrated Human Health Sciences*, *2*(1):1-13.

SSJ PUBLICATIONS
Journal of Integrated Human Health Sciences
ISSN-3143-8814
Research Article

The main hypotheses were that: (1) Mean, SE and Worst measurement domains differ in diagnostic information with the same number and types of underlying nuclear characteristics. (2) the integration of measurement domains will decrease model-dependent prediction uncertainty and (3) non-trivial subset of lesions will show cross-domain diagnostic discordance giving the example that the same biological lesion can produce different reconstructions as we represent it through different quantitative summaries.

## Operational Measurement-Domain Formulation

Let $\beta_i\hat{}(d)$ represent a measurable translation of lesion i in the measurement-summary domain d; d ∈ {Mean, SE, Worst}. The same ten named nuclear characteristics are contained within each domain but a different summary operator.

$\beta_i\hat{}(d)=\{x_i1\hat{}(d), x_i2\hat{}(d), \ldots, x_i10\hat{}(d)\}$

A diagnostic model maps the accessible domain to a malignant probability:

$p_i\hat{}(d) = P(Y_i=\text{malignant} \mid \beta_i\hat{}(d))$

For a binary probability p, model-dependent diagnostic prediction entropy was defined as:

$H_i\hat{}(d)=-p_i\hat{}(d) \log_2 p_i\hat{}(d) - [1 - p_i\hat{}(d)] \log_2[1 - p_i\hat{}(d)]$

A decrease in entropy indicates increased concentration of model probability, but no guarantee of correctness. As such, we did not settle for using entropy just on its own but rather assessed it together with discrimination, accuracy, and Brier score. This is a more general requirement that uncertainty measures need to be evaluated together with predictive performance [7,8].

The reduction in prediction entropy after expanding the accessible domain from A to B was defined as:

$\Delta H_i\ (A \rightarrow B)=H_i\hat{}(A) - H_i\hat{}(B)$

A positive ΔH indicates greater confidence in the increased model for that case. The precise meaning of this number is model dependent and is a prediction entropy reduction; it is never to be construed as a thermodynamic entropy or as a physical property of the lesion.

Cross-domain discordance was defined when the binary classifications produced by Mean, SE, and Worst were not all identical:

$D_i=1$ if $\min_d\ \hat{Y}_i\hat{}(d) \neq \max_d\ \hat{Y}_i\hat{}(d)$; otherwise $D_i=0$

## Materials and Methods

### Dataset

The dataset from Wisconsin Diagnostic Breast Cancer was used for analysis. There are 569 breast mass cases in terms of FNA. Of these, 357 are benign and 212 are malignant (lung cancer). Thirty predictors with real values were computed from imaged cell nuclei. The ten fundamental properties were radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, and fractal dimension. The data set provides the mean, the



**Citation:** Alkhatib A.J (2026). "Diagnostic Information Is Measurement-Domain Dependent Reconstruction Entropy and Cross-Domain Discordance in Breast Fine-Needle Aspiration Cytology A proof-of-concept Operational Measurement-Domain Analysis Using the Wisconsin Diagnostic Breast Cancer Dataset" *Journal of Integrated Human Health Sciences, 2*(1):1-13.

SSJ PUBLICATIONS
Journal of Integrated Human Health Sciences
ISSN-3143-8814
Research Article

standard error, and the "worst" value for each characteristic. The source documentation defines the "worst" value as the mean of the three highest measured values. All predictor values are available.

The uploaded archive contains only the derived tabular measurements and dataset documentation; it does not contain the original cytology images. Consequently, the present analysis concerns diagnostic information in quantitative measurement summaries, not spatial microdomains within the raw image (Table 1).

| Characteristic | Value |
|---|---|
| Cases | 569 |
| Benign | 357 (62.7%) |
| Malignant | 212 (37.3%) |
| Image source | Digitized fine-needle aspiration image of a breast mass |
| Underlying nuclear characteristics | 10 |
| Summary domains | Mean, standard error (SE), Worst |
| Features per single domain | 10 |
| Total features | 30 |
| Missing values | None |
| Raw images in uploaded archive | No |

**Table 1.** Dataset structure and operational measurement domains

## Prediction models and validation

We used a deliberately clear classifier because our aim was to compare information domains, rather than optimize algorithmic complexity. Each domain was modeled with L2-regularized logistic regression after standardization. The same modeling procedure was done on Mean, SE, Worst, their pairwise combinations, and all 30 features.

We used a deliberately clear classifier because our aim was to compare information domains, rather than optimize algorithmic complexity. Each domain was modeled with L2-regularized logistic regression after standardization. The same modeling procedure was done on Mean, SE, Worst, their pairwise combinations, and all 30 features.

There were no efforts to establish external clinical validity from the internal validation. According to Collins and colleagues [9] and Moons and colleagues [10], current recommendations of reporting and risk-of-bias highlight the fact that internal performance cannot replace evaluation in other populations.

## Performance and uncertainty measures

They summarized diagnostic discrimination through area under the receiver operating characteristic curve (AUC). Accuracy, sensitivity and specificity summarises classification at a malignant-probability threshold of 0.50. The Brier score and log loss were used to assess performance. The concentration of probability across the measurement domains is quantified with binary prediction entropy [12,13].

Using 2,500 stratified bootstrap resamples we estimated 95% confidence intervals for AUC and accuracy, mean entropy and entropy reduction; where separate resampling sampling of benign and malignant cases was done with replacement to preserve class structure.



**Citation:** Alkhatib A.J (2026). "Diagnostic Information Is Measurement-Domain Dependent Reconstruction Entropy and Cross-Domain Discordance in Breast Fine-Needle Aspiration Cytology A proof-of-concept Operational Measurement-Domain Analysis Using the Wisconsin Diagnostic Breast Cancer Dataset" *Journal of Integrated Human Health Sciences, 2*(1):1-13.

SSJ PUBLICATIONS
Journal of Integrated Human Health Sciences
ISSN-3143-8814
Research Article

### Cross-domain discordance and paired comparison

This can be reworded as–when Mean, SE, and Worst did not lead to identical binary classifications. The accuracy of these challenging cases were assessed for each single domain and for the 30-feature integrated model. Using exact McNemar tests, we examined classification differences between pairs for Worst versus Mean, Worst versus SE, and for All 30 versus Worst.

### Matched characteristic analysis

A univariate AUC was calculated for each of the ten morphological characteristics with "Mean," "SE," and "Worst" versions to check whether the summary operator changed information by this characteristic with the same name. Because some traits may hold opposite orientations, AUC was represented as max (AUC, 1−AUC), which refers to discrimination magnitude irrespective of its orientation.

## Results

### Diagnostic performance differed substantially by measurement domain

The single-domain that turned out to be the least discriminative was SE with AUC 0.943 and accuracy 87.7%. The Mean values were a lot more informative with AUC 0.985; whereas, the Worst domain had AUC 0.992 and accuracy 97.7 %. The AUC was highest for SE + Worst (0.996), while Mean + Worst and All 30 both had an accuracy of 97.9%. The fact that Worst alone and integrated models are so similar is interesting. The important conclusion is that adding features did not lead to a large gain in conventional classification accuracy, suggesting that Worst summaries already captured most of the diagnostically useful signal (Table 2, Figure 1).

| Domain | Features | AUC (95% CI) | Accuracy | Sensitivity | Specificity | Brier | Entropy |
|---|---|---|---|---|---|---|---|
| Mean | 10 | 0.985 (0.976–0.992) | 93.8% | 89.6% | 96.4% | 0.045 | 0.223 |
| SE | 10 | 0.943 (0.922–0.960) | 87.7% | 77.4% | 93.8% | 0.091 | 0.455 |
| Worst | 10 | 0.992 (0.982–0.999) | 97.7% | 95.8% | 98.9% | 0.020 | 0.162 |
| Mean + SE | 20 | 0.986 (0.978–0.994) | 95.8% | 92.9% | 97.5% | 0.037 | 0.190 |
| Mean + Worst | 20 | 0.992 (0.982–0.999) | 97.9% | 95.8% | 99.2% | 0.020 | 0.144 |
| SE + Worst | 20 | 0.996 (0.990–0.999) | 97.7% | 95.3% | 99.2% | 0.020 | 0.137 |
| All 30 | 30 | 0.995 (0.989–0.999) | 97.9% | 96.7% | 98.6% | 0.020 | 0.125 |

**Table 2.** Repeated cross-validated performance by accessible measurement domain



**Citation:** Alkhatib A.J (2026). "Diagnostic Information Is Measurement-Domain Dependent Reconstruction Entropy and Cross-Domain Discordance in Breast Fine-Needle Aspiration Cytology A proof-of-concept Operational Measurement-Domain Analysis Using the Wisconsin Diagnostic Breast Cancer Dataset" *Journal of Integrated Human Health Sciences*, *2*(1):1-13.

SSJ PUBLICATIONS
Journal of Integrated Human Health Sciences
ISSN-3143-8814
Research Article

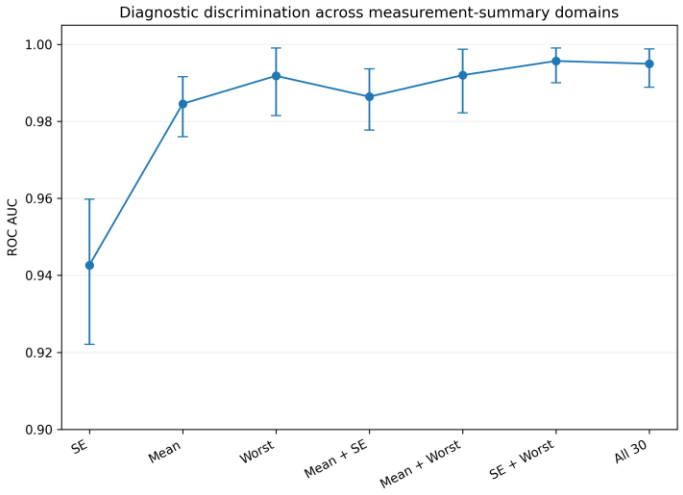



**Figure 1.** Diagnostic discrimination across measurement-summary domains. Points show AUC calculated from averaged repeated out-of-fold probabilities; error bars show stratified bootstrap 95% confidence intervals

## Prediction entropy was domain dependent

In terms of mean binary prediction entropy, the SE domain (0.455 bits) had the highest value, followed by Mean (0.223), Worst (0.162), and the integrated All-30 model (0.125). As a result, varying quantitative summaries of the same lesion not only resulted in different discrimination but also indifferent levels of model confidence (Figure 2).

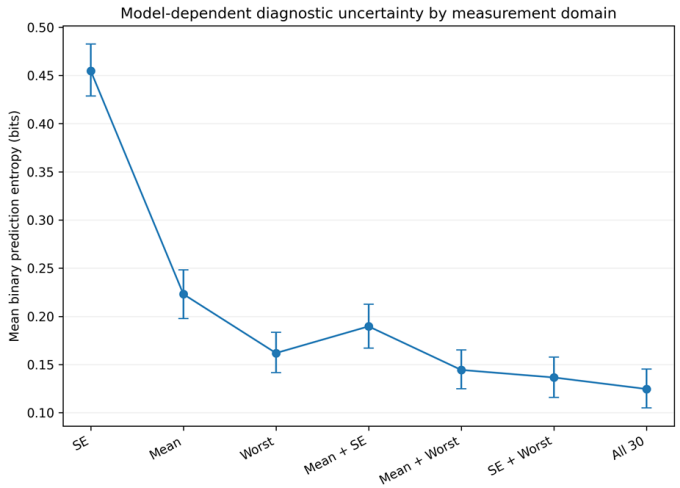



**Figure 2.** Mean model-dependent binary prediction entropy by measurement domain. Lower values indicate more concentrated malignant/benign probabilities, not necessarily greater biological certainty

## Integration produced the largest uncertainty reduction when the starting domain was weak

The use of SE across all 30 features led to a mean reduction in prediction entropy by 0.330 bits in the models, with 92.1 per cent of lesions showing a reduction in entropy. Mean produced a smaller reduction of 0.098 bits. Initiating from Worst generated a mean reduction of only 0.037 bits, consistent with the highly effective Worst domain (Table 3, Figure 3).



**Citation:** Alkhatib A.J (2026). "Diagnostic Information Is Measurement-Domain Dependent Reconstruction Entropy and Cross-Domain Discordance in Breast Fine-Needle Aspiration Cytology A proof-of-concept Operational Measurement-Domain Analysis Using the Wisconsin Diagnostic Breast Cancer Dataset" *Journal of Integrated Human Health Sciences*, *2*(1):1-13.

SSJ PUBLICATIONS
Journal of Integrated Human Health Sciences
ISSN-3143-8814
Research Article

| From | To | Mean entropy reduction (bits) | Cases with lower entropy (%) | Delta AUC | Delta accuracy |
|---|---|---|---|---|---|
| SE | All 30 | 0.330 (0.304–0.356) | 92.1% | +0.052 | +10.2 pp |
| Mean | All 30 | 0.098 (0.080–0.117) | 86.8% | +0.010 | +4.0 pp |
| Worst | All 30 | 0.037 (0.029–0.046) | 88.0% | +0.003 | +0.2 pp |
| SE | SE + Worst | 0.318 (0.292–0.346) | 91.6% | +0.053 | +10.0 pp |
| Mean | Mean + Worst | 0.079 (0.063–0.094) | 85.1% | +0.007 | +4.0 pp |

**Table 3.** Incremental prediction-entropy reduction after expanding the accessible measurement domain

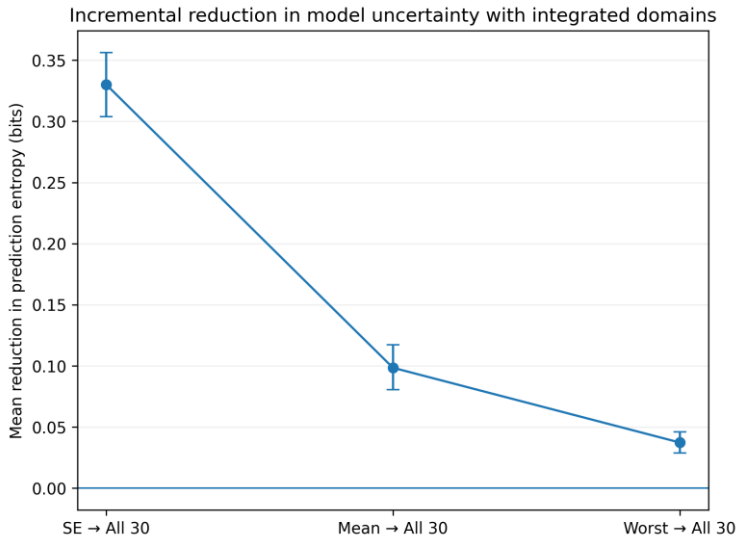



**Figure 3.** Mean reduction in model prediction entropy when single domains are expanded to the integrated 30-feature representation. Error bars show stratified bootstrap 95% confidence intervals

## Nearly 15% of lesions were classified differently across single domains

There was a disagreement in classification of 85 out of 569 (14.9%) lesions between Mean, SE, and Worst. In the other 484 consensus cases, the agreement among the three single domains and correct classification was 99.2% (Table 4).

| Domain | Accuracy among discordant cases | Correct | Discordant N |
|---|---|---|---|
| Mean | 63.5% | 54 | 85 |
| SE | 22.4% | 19 | 85 |
| Worst | 89.4% | 76 | 85 |
| All 30 | 91.8% | 78 | 85 |

**Table 4.** Performance restricted to cross-domain discordant cases

When it comes to discordant lesions, the Mean domain was the most accurate at 63.5% of the time, followed by the SE domain at 22.4% of the time, the Worst domain at 89.4% of the time, and finally the integrated All-30 model was correct in 91.8% of the cases. Thus, most of the practical difference among the domains were these difficult cases (Figure 4).



**Citation:** Alkhatib A.J (2026). "Diagnostic Information Is Measurement-Domain Dependent Reconstruction Entropy and Cross-Domain Discordance in Breast Fine-Needle Aspiration Cytology A proof-of-concept Operational Measurement-Domain Analysis Using the Wisconsin Diagnostic Breast Cancer Dataset" *Journal of Integrated Human Health Sciences*, *2*(1):1-13.

SSJ PUBLICATIONS
Journal of Integrated Human Health Sciences
ISSN-3143-8814
Research Article

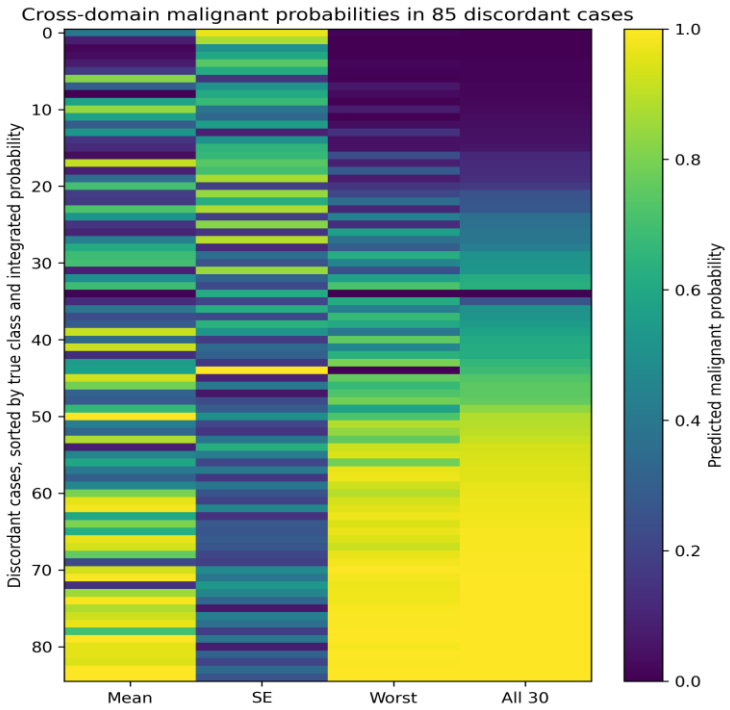



**Figure 4.** Predicted malignant probabilities for the 85 lesions in which Mean, SE, and Worst single-domain classifications disagreed. Rows are sorted by true class and integrated-model probability. The 0.50 threshold separates benign from malignant classification.

## Worst outperformed Mean and SE in paired decisions

The Worst classification correctly identified 27 lesions that the Mean model misclassified. The Mean model, on the other hand, correctly classified only 5 lesions that the Worst classification misidentified (exact McNemar p=0.000113). Compared to SE, Worst was significantly more accurate. There were 65 instances in which SE was incorrect, but Worst was correct. In 8 instances, SE was correct, but Worst was incorrect (p<0.001). In contrast, there was no significant paired classification difference for All 30 and Worst (Table 5).

| Model A | Model B | A correct/B wrong | A wrong/B correct | Exact McNemar p |
|---|---|---|---|---|
| Worst | Mean | 27 | 5 | <0.001 |
| Worst | SE | 65 | 8 | <0.001 |
| All 30 | Worst | 6 | 5 | 1.000 |

**Table 5.** Exact paired comparison of lesion-level classification errors

## The effect of the summary operator was visible feature by feature

The influence of the summary-domain was beyond that. The worst measurements of perimeter, radius, area and concave-point showed very high univariate discrimination, whereas several corresponding measures of SE were much weaker. An example of this is texture had AUC 0.776 as Mean summary, 0.512 as SE, and 0.785 as Worst; smoothness had AUC 0.722, 0.531, and 0.754 as a summary (Table 6, Figure 5).



**Citation:** Alkhatib A.J (2026). "Diagnostic Information Is Measurement-Domain Dependent Reconstruction Entropy and Cross-Domain Discordance in Breast Fine-Needle Aspiration Cytology A proof-of-concept Operational Measurement-Domain Analysis Using the Wisconsin Diagnostic Breast Cancer Dataset" *Journal of Integrated Human Health Sciences*, *2*(1):1-13.

SSJ PUBLICATIONS
Journal of Integrated Human Health Sciences
ISSN-3143-8814
Research Article

| Feature | Mean | SE | Worst |
|---|---|---|---|
| Area | 0.938 | 0.926 | 0.970 |
| Compactness | 0.864 | 0.727 | 0.862 |
| Concave Points | 0.964 | 0.792 | 0.967 |
| Concavity | 0.938 | 0.781 | 0.921 |
| Fractal Dimension | 0.515 | 0.620 | 0.686 |
| Perimeter | 0.947 | 0.876 | 0.975 |
| Radius | 0.938 | 0.868 | 0.970 |
| Smoothness | 0.722 | 0.531 | 0.754 |
| Symmetry | 0.699 | 0.555 | 0.737 |
| Texture | 0.776 | 0.512 | 0.785 |

**Table 6.** Univariate AUC for matched nuclear characteristics across summary domains

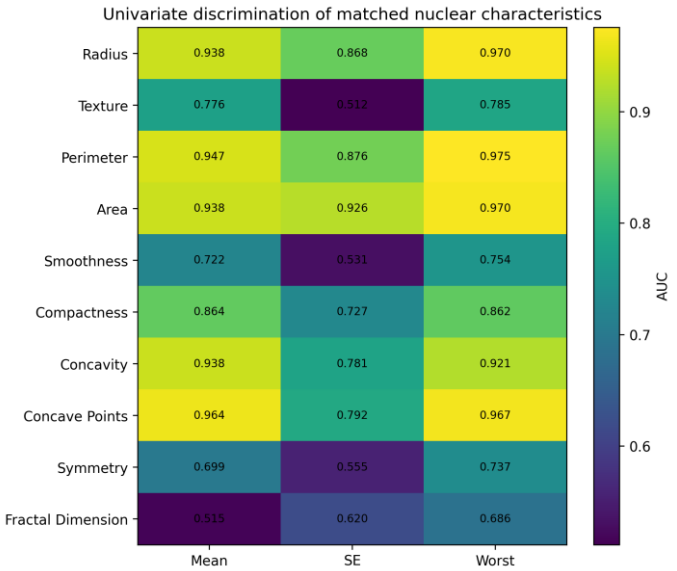



**Figure 5.** Univariate discrimination of the same ten nuclear characteristics under Mean, SE, and Worst summary operators. Each cell displays AUC magnitude

## Discussion

The article revolves around a not-so-big but important difference the biological specimen and the quantitative measure domain are not the same using a classical breast-cytology dataset. Diagnostic information from Mean, SE, and Worst summaries was quite different although they both were derived from the same set of ten named nuclear characteristics contained in digitized FNA images.

The Worst summaries received the most powerful single-domain result. It makes biological sense because in the presence of a malignant cytology there is nuclear enlargement, nuclear pleomorphism, contour irregularity, and focal extremes which may dilute through a simple averaging. The original WDBC work identified nuclear size and associated morphological features to be very informative, and combinations of Worst Area, Worst Smoothness, and Mean Texture were reported to provide great discrimination [3-5,14]. This study builds on earlier research that looked at complete matched domains and assessed probability uncertainty and cross-domain disagreement.



**Citation:** Alkhatib A.J (2026). "Diagnostic Information Is Measurement-Domain Dependent Reconstruction Entropy and Cross-Domain Discordance in Breast Fine-Needle Aspiration Cytology A proof-of-concept Operational Measurement-Domain Analysis Using the Wisconsin Diagnostic Breast Cancer Dataset" *Journal of Integrated Human Health Sciences*, *2*(1):1-13.

SSJ PUBLICATIONS
Journal of Integrated Human Health Sciences
ISSN-3143-8814
Research Article

The SE domain was always the weakest. This does not mean that variability is not biologically important. The particular SE summaries found in WDBC were, however, less discriminative under the present model, than the Mean and Worst summaries. This outcome shows that the informational worth of a feature can't be ascertained just by its biological label as these variables summarize the same source image differently and the measurement operator matters. Disagreement across domains was especially revealing. About 15% of the lesions crossed the classification threshold 0.50 differently by Mean, SE, or Worst information. The much larger consensus group had an accuracy over 99%. Most of the remaining errors and domain differences were in the discordant group. So, like one can use the concept of practical uncertainty, where conflicting measures with legitimate representation of the same source, indicate a case worth further review, or further information.

By combining the domains, the entropy of prediction is decreased however that depends strongly on the representation from which one is starting. When all domains were added to SE, there was a considerable entropy reduction, while Worst only produced a slight reduction. This coincides with lessening returns of information once a discretely identifiable representation is already present. 'More variables' and 'more information' are clearly not the same thing, as the above result shows. A few carefully chosen measurements can capture most of the relevant signal, and adding many correlated measurements would mainly refine confidence rather than help classification.

The prediction entropy of this study must be interpreted with caution. A prediction of low entropy may still be wrong and the concentration of probability depends on model specification and calibration. Thus, entropy was analyzed together with AUC, accuracy of classification, Brier score and paired errors. The methodology of clinical prediction is increasingly focused on calibration, uncertainty and applicability rather than merely discrimination [7-9].

One interpretation of the analysis is as an empirical biomedical equivalent of an operational reconstruction. Let β the information obtained from a biological specimen under a defined measurement protocol. Various β may aid diverse reconstructions of one similar subclinical state. This does not indicate that measuring an illness creates the state or lesion(s) are multivalent or polymorphous. Inference depends on the representation that is available. In biomedical terminology, the important distinction is between the disease state, measurement state and reconstructed diagnostic state.

This differentiation is progressively significant in modern computational pathology. Artificial intelligence systems now work with whole slide images, cytology, radiology, molecular data, and multimodal combinations to create multiple views of the same patient. Techniques in Biomedical Citation Analysis for Cancer-AI and Cancer Cytopathology Reviews emphasize recognizing the potential of computation analysis and the need for robust and validated standardization and clinically meaningful interpretation [11,15,16]. The present evidence suggests that domain-specific information and domain discordance should be evaluated explicitly when evaluating multimodal or multiview diagnostic systems.

The principle question was sufficiently simple so that highly complex machine-learning algorithm were not necessary, and the study shows this methodologically. Transparent probabilities from logistic regression can perform just as well as more complex methods in clinical prediction studies if properly validated [17]. The comparison of measurement domains and the structure of uncertainty are new, rather than simply maximizing benchmark accuracy.



**Citation:** Alkhatib A.J (2026). "Diagnostic Information Is Measurement-Domain Dependent Reconstruction Entropy and Cross-Domain Discordance in Breast Fine-Needle Aspiration Cytology A proof-of-concept Operational Measurement-Domain Analysis Using the Wisconsin Diagnostic Breast Cancer Dataset" *Journal of Integrated Human Health Sciences, 2*(1):1-13.

SSJ PUBLICATIONS
Journal of Integrated Human Health Sciences
ISSN-3143-8814
Research Article

## Proposed Biomedical Principles

The findings support three cautious methodological propositions that can be tested in other biomedical datasets:

1. Measurement-Domain Dependence: diagnostic information is conditioned by the representation through which the biological specimen is accessed.

$$I(Y; \beta^{(a)}) \text{ need not equal } I(Y; \beta^{(b)})$$

2. Reconstruction Non-Uniqueness: the same lesion can receive different probabilistic reconstructions when different legitimate measurement domains are used.

$$P(Y \mid \beta^{(a)}) \text{ need not equal } P(Y \mid \beta^{(b)})$$

3. Integration Gain Is Context Dependent: adding a new measurement domain can reduce uncertainty substantially when the starting domain is weak, but may add little when the existing domain already captures most discriminative information.

$$\Delta H(A \rightarrow A+B) \text{ depends on the information already present in } A$$

These propositions describe inference and measurement; they are not claims about the fundamental biological ontology of cancer.

## Limitations

- The previously used dataset with one source population having 569 cases cannot be externally validated in a contemporary and geographically diverse cohort.
- The uploaded archive contains derived nuclear measurements, not the raw cytology images so the spatial tumor heterogeneity and true image microdomains could not be assessed.
- Mean, SE and Worst are correlated summaries derived from the same nuclei and should not be interpreted as independent biological compartments.
- The "Worst" value in WDBC is summary value which is the mean of three largest measurements & not single maximums.
- Repeated cross-validation allows for robust internal comparisons, but fails to establish transportability to current clinical practice.
- Prediction entropy which is not biological or thermodynamic entropy is a model-reliant entropy.
- The 0.50 decision threshold was used for comparability, not because it represents an optimized or clinically validated threshold.
- A single transparent modeling family has been used. The probability estimates may be changed by alternative algorithms, but the measurement-domain comparison at the center should stay testable.
- The FNAC practice, breast diagnostic pathways, and AI technology development have evolved remarkably since the original WDBC data was collected.


**Citation:** Alkhatib A.J (2026). "Diagnostic Information Is Measurement-Domain Dependent Reconstruction Entropy and Cross-Domain Discordance in Breast Fine-Needle Aspiration Cytology A proof-of-concept Operational Measurement-Domain Analysis Using the Wisconsin Diagnostic Breast Cancer Dataset" *Journal of Integrated Human Health Sciences, 2*(1):1-13.

SSJ PUBLICATIONS
Journal of Integrated Human Health Sciences
ISSN-3143-8814
Research Article

## Conclusion

The diagnostic information in the Wisconsin Diagnostic Breast Cancer dataset depends heavily on how the same nuclear morphology was summarized. The SE domain, despite having ten versions of the same named morphological characteristics, was outperformed significantly by the Worst feature domain. Combining domains lowered model prediction entropy, especially when starting from a weak representation, and nearly 15% of lesions disagreed on Mean/SE/Worst classifications. The findings enable a solid biomedical proof of concept that can separate the biological specimen from its operational measurement representation. The outcome shows that the lesion does not alter, but the diagnostic reconstruction does, depending on the information of the measurement protocol. The framework should be tested on more realistic spatial histopathology patches, multimodal imaging-molecular datasets and repeated measures where access domains correspond to physically distinct biological regions or modalities.

## Data availability

The study used the publicly available Wisconsin Diagnostic Breast Cancer dataset. The supplied archive contains wdbc.data and wdbc.names and contains no raw cytology images.

## Ethics statement

This study is a secondary analysis of a publicly released de-identified dataset and involved no new participant recruitment or intervention.

## References

1. Paul, P., Azad, S., Agrawal, S., Rao, S., & Chowdhury, N. (2023). Systematic review and meta-analysis of the diagnostic accuracy of the International Academy of Cytology Yokohama System for reporting breast fine-needle aspiration biopsy in diagnosing breast cancer. *Acta cytologica*, *67*(1), 1-16.
2. Schmitt, F. C. (2024). Overcoming Pitfalls in Breast FNAC: A Practical Review. *Acta Cytologica*.
3. Wolberg, W. H., Street, W. N., & Mangasarian, O. L. (1994). Machine learning techniques to diagnose breast cancer from image-processed nuclear features of fine needle aspirates. *Cancer letters*, *77*(2-3), 163-171.
4. Street, N. (1995). Image analysis and machine learning applied to breast cancer diagnosis and prognosis. *Analytical and quantitative cytology and histology/the International Academy of Cytology [and] American Society of Cytology*.
5. Wolberg, W. H., Street, W. N., Heisey, D. M., & Mangasarian, O. L. (1995). Computerized breast cancer diagnosis and prognosis from fine-needle aspirates. *Archives of Surgery*, *130*(5), 511-516.
6. Wolberg, M. W. H., Street, M. W. N., Heisey, P. D. M., & Mangasarian, P. O. L. (1995). Computer-derived nuclear features distinguish malignant from benign breast cytology. *Human pathology*, *26*(7), 792-796.
7. Van Calster, B., McLernon, D. J., Van Smeden, M., Wynants, L., Steyerberg, E. W., & Topic Group 'Evaluating diagnostic tests and prediction models' of the STRATOS initiative Bossuyt Patrick Collins Gary S. Macaskill Petra McLernon David J. Moons Karel GM Steyerberg Ewout W. Van Calster Ben van Smeden Maarten Vickers Andrew J. (2019). Calibration: the Achilles heel of predictive analytics. *BMC medicine*, *17*(1), 230.



**Citation:** Alkhatib A.J (2026). "Diagnostic Information Is Measurement-Domain Dependent Reconstruction Entropy and Cross-Domain Discordance in Breast Fine-Needle Aspiration Cytology A proof-of-concept Operational Measurement-Domain Analysis Using the Wisconsin Diagnostic Breast Cancer Dataset" *Journal of Integrated Human Health Sciences*, *2*(1):1-13.

SSJ PUBLICATIONS

Journal of Integrated Human Health Sciences

ISSN-3143-8814

Research Article

8. Chua, M., Kim, D., Choi, J., Lee, N. G., Deshpande, V., Schwab, J., ... & Do, S. (2023). Tackling prediction uncertainty in machine learning for healthcare. *Nature Biomedical Engineering*, *7*(6), 711-718.
9. Collins, G. S., Moons, K. G., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., ... & Logullo, P. (2024). TRIPOD+ AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *bmj*, *385*.
10. Moons, K. G., Damen, J. A., Kaul, T., Hooft, L., Navarro, C. A., Dhiman, P., ... & Van Smeden, M. (2025). PROBAST+ AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *bmj*, *388*.
11. Hays, P. (2024). Artificial intelligence in cytopathological applications for cancer: a review of accuracy and analytic validity. *European Journal of Medical Research*, *29*(1), 553.
12. Glenn, W. B. (1950). Verification of forecasts expressed in terms of probability. *Monthly weather review*, *78*(1), 1-3.
13. Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, *27*(3), 379-423.
14. Wolberg, W., Mangasarian, O., Street, N., & Street, W. (1995). Breast cancer wisconsin (diagnostic). *UCI machine learning repository*, *10*, C5DW2B.
15. Díaz, O., Rodríguez-Ruíz, A., & Sechopoulos, I. (2024). Artificial Intelligence for breast cancer detection: Technology, challenges, and prospects. *European journal of radiology*, *175*, 111457.
16. Katayama, A., Aoki, Y., Watanabe, Y., Horiguchi, J., Rakha, E. A., & Oyama, T. (2024). Current status and prospects of artificial intelligence in breast cancer pathology: convolutional neural networks to prospective Vision Transformers. *International Journal of Clinical Oncology*, *29*(11), 1648-1668.
17. Evangelia, C., Jie, M., Collins, G., Steyerberg, E., Verbakel, J., & Van Calster, B. (2019). A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of Clinical Epidemiology*, *110*.



**Citation:** Alkhatib A.J (2026). "Diagnostic Information Is Measurement-Domain Dependent Reconstruction Entropy and Cross-Domain Discordance in Breast Fine-Needle Aspiration Cytology A proof-of-concept Operational Measurement-Domain Analysis Using the Wisconsin Diagnostic Breast Cancer Dataset" *Journal of Integrated Human Health Sciences*, *2*(1):1-13.