

Authors: Daerin Hwang, Abdel Badih El Ariss, Norawit Kijpaisalratana, Paul Chong, David Chen, MD, Ahmad Hassan, Pedram Safari

Type: Original Research

Keywords: Thoracic Aortic Dissection, Shannon Entropy, Decision Tree Analysis, Entropy Reduction, Diagnostic Uncertainty, Emergency Medicine, Diagnostic Imaging, Information Theory

Edited By: Nicole Joanna Poritsanos

Received: 2025-08-01

Published: 2026-01-22

An Entropy-Based Approach to Decision Tree Analysis in Emergency Medicine: Optimization of Diagnostic Tool Selection to Thoracic Aortic Dissection

Abstract

Introduction

Shannon entropy is a key concept in both machine learning and information theory, significantly influencing decision tree modeling. In emergency medicine, the utilization of testing and imaging tools to reduce uncertainty is vital for enhancing medical decision-making, especially in time-sensitive scenarios. While sensitivity and specificity assess test accuracy, entropy provides insight into how much a test clarifies the overall clinical picture, which is crucial in time-sensitive situations like thoracic aortic dissection. The aim of this study is to evaluate and compare the effectiveness of entropy and entropy reduction against conventional metrics of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), in the diagnostic evaluation of thoracic aortic dissection.

Methodology

We gathered diagnostic metrics, including true positives and negatives, as well as false positives and negatives, from a well-established online diagnostic accuracy database known as **"Get the Diagnosis"** for thoracic aortic dissection. Data collection took place from November 17, 2022, to January 22, 2023. This dataset allowed us to compute sensitivities, specificities, negative predictive values (NPVs), and positive predictive values (PPVs). A decision tree representing the diagnostic tool was created and examined for Shannon entropy and entropy reduction across the various child nodes. The decision tree was based on 2x2 diagnostic tables, covering total cases (N), true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN).

Results

Using a helical CT scan for thoracic aortic dissection, we observed a sensitivity of 97.64% and a specificity of 98.90%, whereas MRI exhibited a lower sensitivity (93.33%) but a higher specificity (99.30%). However, entropy removal calculations revealed that the helical CT scan removed 87.23% of entropy associated with thoracic aortic dissection, while MRI removed 81.36%. Within these populations studied, this suggests helical CT scan may provide greater reduction in diagnostic uncertainty than MRI in thoracic aortic dissection assessment and underscores the unique insights offered by entropy removal, which sensitivity and specificity fail to provide.

Conclusion

Unlike sensitivity and specificity, entropy captures the inherent variability and uncertainty in diagnostic data, especially in complex medical scenarios where it accommodates individual variability in patient responses and disease manifestations.

Introduction

Evidence-based medicine is pivotal in guiding clinical decision-making processes. Its foundation rests on using metrics such as sensitivity and specificity to identify the most suitable diagnostic tests for specific diseases (1). However, the application of these metrics frequently encounters limitations in emergency medicine, thus, we aim to develop dynamic models that can be recalculated but also be used in the future in dynamic situations such as multiple order models.

1. In contrast, the concept of entropy, derived from information theory (3), offers a novel approach to selecting diagnostic tests. Shannon entropy, a principle introduced by Claude Shannon in information theory, quantifies the uncertainty of the predicted probability distribution (disease vs. no disease) at a given decision point (4). This concept accounts for the uncertainty present in a situation and the reduction of this uncertainty through the acquisition of information, a process known as entropy removal. Shannon entropy aids in forecasting the range and variability of outcomes, underscoring the importance of

considering all possible scenarios during decision-making. In emergency medicine, physicians routinely grapple with uncertain situations, and the lack of a robust method to quantify this uncertainty poses a significant obstacle. This becomes especially crucial in situations involving severe illnesses like aortic dissection, which have high mortality rates of 1% to 2% per hour, thereby demanding precise and prompt decision-making (5).

The objective of this study is to assess and compare the efficacy of entropy and entropy reduction with traditional sensitivity and specificity metrics in the diagnostic evaluation of thoracic aortic dissection. The study aims to establish whether entropy-based methods can provide a comparable understanding of uncertainty in diagnosis, potentially leading to improved methodologies for understanding time-sensitive clinical scenarios like those encountered in emergency medicine.

Methodology

IRB statement

This study is exempt from IRB review of Massachusetts General Hospital and Harvard Medical School as research involves collecting and studying existing data of which sources are publicly available, and subjects cannot be identified directly or through identifiers linked to the subjects.

Entropy equations

Shannon entropy is a mathematical method used to calculate the expected uncertainty of a probability distribution. For a binary diagnostic outcome, Shannon entropy H is defined as stated in Eq. (1):

$$H = - \sum p \log (p) ,$$

$$i = 1, 2, \dots, n$$

where p 's are the probabilities of each possible outcome of an event. Throughout, we interpret H as the uncertainty in the diagnostic probability distribution (rather than the inherent unpredictability of patient outcomes) and entropy removal as the fractional reduction in that uncertainty after a diagnostic test.

Decision trees consist of primary nodes that divide into subsidiary nodes. The formation of these nodes, and the decisions leading to their division, can be used to create decision tree diagrams for tools in medical diagnosis. This process involves the application of diagnostic values derived from 2x2 diagnostic matrices. In such matrices, 'N' is used to denote the overall sample size of the study, 'TP' signifies the total of true positive results, 'FP' represents the instances of false positives, 'FN' is used for the count of false negatives, and 'TN' indicates the total true negative outcomes.

We compiled the metrics for medical imaging techniques, clinical symptoms, and signs, as well as laboratory tests utilized in diagnosing aortic dissection, as outlined in Table 1. For each specific finding or test, we selected one random study that calculated the metrics including true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN)(7-14). 'Get the Diagnosis' was used to identify the underlying peer-reviewed studies from which we extracted the TP, FP, FN, and TN values; one representative published study was selected for each diagnostic tool. This pragmatic approach allowed us to compute entropy-based metrics across multiple diagnostic modalities.

Diagnostic Tool	Type	TP	FP	FN	TN
Imaging					
MRI	Imaging	644	19	46	2693
Helical CT	Imaging	207	8	5	719
TEE	Imaging	1258	40	69	5509
CXR	Imaging	38	47	19	112
Non-contrast CT	Imaging	77	5	8	80
Laboratory Test					
D-dimer in high risk	Laboratory test	79	78	2	46
D-dimer in intermediate risk	Laboratory test	131	225	2	150
D-dimer in low risk	Laboratory test	19	211	0	90
Acute renal failure	Laboratory Test	20	108	0	122
Leukocyte count $\geq 15 \times 10^9/L$	Laboratory test	33	95	27	95

Symptom					
Immediate onset of pain	Symptom	101	27	37	85
Trigger of pain	Symptom	19	109	28	94
Prodromal symptoms	Symptom	13	115	15	107
Intense severity of pain	Symptom	110	18	67	55
Tearing or ripping pain	Symptom	79	49	7	115
Migratory pain	Symptom	56	72	7	115
Pleuritic pain	Symptom	12	116	21	101
Posterior chest or lower back pain	Symptom	64	64	31	91
Pain in neck or jaw and/or $\geq$ 1 extremity	Symptom	34	94	14	108
Abdominal pain	Symptom	28	100	14	108
Anterior chest pain	Symptom	97	31	85	37

Focal neurologic signs	Symptom	17	111	0	122
Syncope	Symptom	13	115	12	110
Prolonged loss of consciousness or coma	Symptom	3	125	2	120
Myocardial infarction syndrome	Symptom	3	125	2	120
Physical exam					
Pulse and/or blood pressure differentials	Physical exam	49	79	1	121
Murmur of aortic regurgitation	Physical exam	51	77	29	93
Hypertension_ on_emergency department admission	Physical exam	53	75	38	84
Hypotension_ on_emergency department admission	Physical exam	15	113	7	115
Abdominal signs	Physical exam	13	115	14	108

Findings					
Mediastinal and/or aortic widening	Imaging finding	97	31	27	95
Left ventricular hypertrophy	Imaging finding	27	101	8	114
Pleural effusion	Imaging finding	17	111	22	100
Previous Q-wave infarction	EKG finding	9	119	15	107

Table 1: TP, FP, TN and FN for different diagnostic tools derived from peer reviewed manuscripts.

Using diagnostic metrics (N, TP, FP, etc.), we applied the entropy calculation approach described by He et al. (6) to calculate Entropy removal for each medical diagnostics. For each diagnostic test, the values for p were derived from the 2 x 2 table of TP, FP, FN, and TN. For instance, the i pre-test probabilities of disease and no disease were computed as $(TP + FN)/N$ and $(FP+TN)/N$, respectively, where N is the total sample size.

Entropy removal for a test was defined as

$$(H - H')$$

$$\text{Entropy removal} = \text{before} - \text{after}, (2)$$

H

before

where

H'

before

is the entropy before applying the test (parent node) and H' is the weighted *after* average entropy of the post-test states (children nodes) based on the proportions of positive and negative test results.

For each diagnostic test, we utilized the TP, FP, FN, and TN values as reported in a single published study, reflecting the patient selection and disease prevalence of that particular cohort. For this reason, the entropy and corresponding entropy reduction values are inherently context-specific and rely on the underlying pre-test probability for each study. Thus, entropy values across

different diagnostics must be evaluated as illustrative within their respective study populations rather than as directly generalizable.

Results

The entropy reduction, calculated from sensitivity, specificity, PPV, and NPV for each diagnostic tool, is summarized in Table 2.

Diagnostic Tool	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Removed entropy (%)
Imaging					
Helical CT (clinician)	97.64%	98.90%	96.28%	99.31%	87.23%
TEE (clinician)	94.80%	99.28%	96.92%	98.76%	83.69%
MRI (clinician)	93.33%	99.30%	97.13%	98.32%	81.36%
Non-contrast CT (clinician)	90.59%	94.12%	93.90%	90.91%	61.27%
CXR (clinician)	66.67%	70.44%	44.71%	85.50%	9.61%
Mediastinal and/or aortic widening	78.23%	75.40%	75.78%	77.87%	21.89%
Left ventricular hypertrophy	77.14%	53.02%	21.09%	93.44%	5.70%
Pleural effusion	43.59%	47.39%	13.28%	81.97%	0.50%
Laboratory Tests					

D-dimer in low risk	100.00%	29.90%	8.26%	100.00%	9.03%
D-dimer in intermediate risk	98.50%	40.00%	36.80%	98.68%	16.17%
D-dimer in high risk	97.53%	37.10%	50.32%	95.83%	14.84%
Acute renal failure	100.00%	53.04%	15.63%	100.00%	20.40%
Leukocyte count $\geq 15 \times 10^9/L$	55.00%	50.00%	25.78%	77.87%	0.17%
Physical Examination					
Pulse and/or blood pressure differentials	98.00%	60.50%	38.28%	99.18%	27.28%
Murmur of aortic regurgitation	63.75%	54.71%	39.84%	76.23%	2.39%
Hypertension on emergency department admission	58.24%	52.83%	41.41%	68.85%	0.87%
Hypotension on emergency department admission	68.18%	50.44%	11.72%	94.26%	1.91%

Abdominal pain	66.67%	51.92%	21.88%	88.52%	2.17%
Symptoms					
Tearing or ripping pain	91.86%	70.12%	61.72%	94.26%	30.41%
Focal neurologic signs	100.00%	52.36%	13.28%	100.00%	19.27%
Migratory pain	88.89%	61.50%	43.75%	94.26%	18.85%
Immediate onset of pain	73.19%	75.89%	78.91%	69.67%	18.10%
Intense severity of pain	62.15%	75.34%	85.94%	45.08%	9.95%
Posterior chest or lower back pain	67.37%	58.71%	50.00%	74.59%	4.91%
Pain in neck or jaw and/or $\geq$ 1 extremity	70.83%	53.47%	26.56%	88.52%	3.85%
Pleuritic pain	36.36%	46.54%	9.38%	82.79%	1.73%
Trigger of pain	40.43%	46.31%	14.84%	77.05%	1.12%
Anterior chest pain	53.30%	54.41%	75.78%	30.33%	0.40%
Prolonged loss of	60.00%	48.98%	2.34%	98.36%	0.33%

consciousness or coma					
Myocardial infarction syndrome	60.00%	48.98%	2.34%	98.36%	0.33%
Prodromal symptoms	46.43%	48.20%	10.16%	87.70%	0.16%
Abdominal signs	48.15%	48.43%	10.16%	88.52%	0.07%
Syncope	52.00%	48.89%	10.16%	90.16%	0.00%
ECG Findings					
Previous Q-wave infarction	37.50%	47.35%	7.03%	87.70%	1.27%

Table 2: Comparison of Sensitivity, specificity, negative predictive value, positive predictive value and entropy reduction

The table provides a comprehensive overview of various diagnostic tools and clinical observations, quantifying their effectiveness using statistical measures like sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and importantly, entropy reduction.

Entropy reduction can be used as a measure of the effectiveness of a diagnostic tool in reducing uncertainty in diagnosis. High entropy reduction suggests that the test clarifies the diagnostic picture by reducing diagnostic uncertainty, which is crucial in clinical decision-making.

For instance, both MRI and helical CT scans exhibit high sensitivity and specificity in diagnosing aortic dissection. However, upon analyzing entropy removal, helical CT demonstrates a superior value compared to MRI within the study population, as illustrated in Table 2. In contrast, a test like D-dimer in high-risk patients, despite its high sensitivity of 97.53%, has a relatively lower specificity of 37.1% and PPV of 50.32%, contributing to its lower removed entropy of 14.84%.

Similarly, clinical observations such as immediate onset of pain, intense severity of pain, and prolonged loss of consciousness or coma, achieved entropy reduction of only 18.10%, 9.95%, and 0.33%, respectively. These clinical signs, therefore, provide limited effectiveness in reducing diagnostic uncertainty compared to advanced diagnostic tools like MRI. The low entropy reduction values suggest that relying solely on symptoms can leave a significant degree of uncertainty in diagnosis.

Overall, the removed entropy metric serves as a critical indicator of how much a diagnostic tool or clinical observation contributes to reducing uncertainty in diagnosis. Higher values in this metric are particularly desirable in clinical settings, as they indicate a more definitive contribution to diagnosing or ruling out a condition.

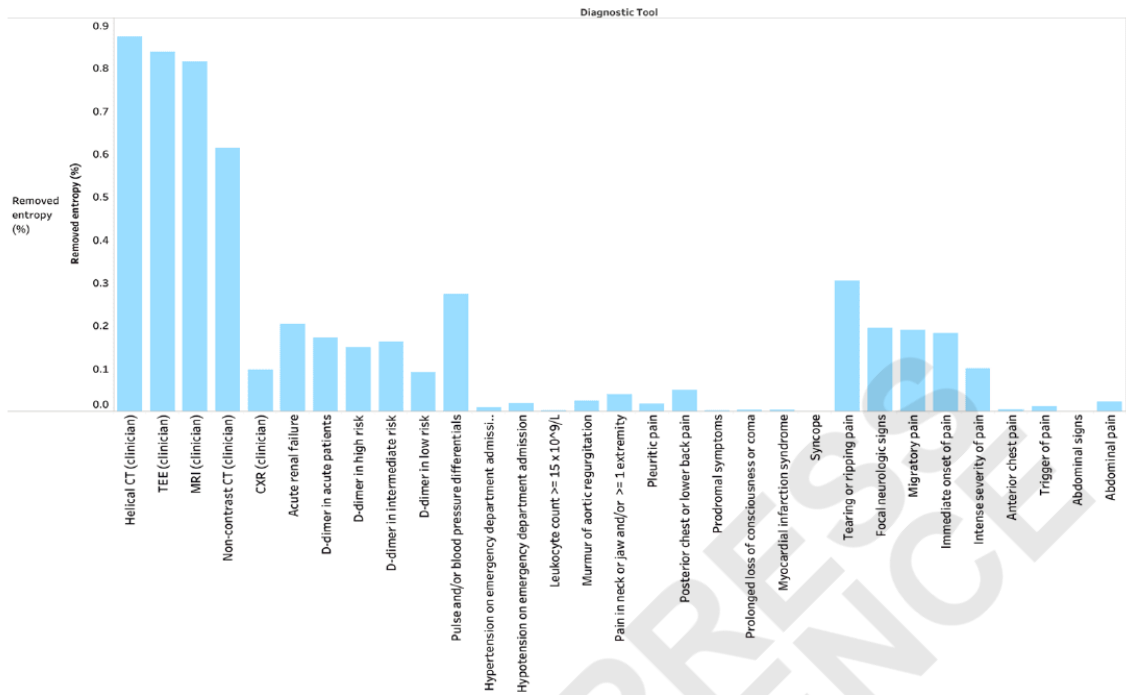


Figure 1: Entropy reduction in percentage

Figure 1: Entropy reduction in percentage

Among the imaging modalities, helical CT had the highest entropy reduction at 87.23%. Mediastinal widening had the highest entropy (21.89%) when comparing it to other clinical findings. Occurrence of acute renal failure (20.41%) and the experience of tearing or ripping pain (30.41%) demonstrated a higher reduction in diagnostic uncertainty compared to other laboratory findings and types of pain reported, respectively. Figure 1 preserves its original fixed ordering, while Table 2 presents the fully grouped and sorted entropy-removal values.

Additionally, by examining Figure 1 from high to low reduced entropy (left to right), it becomes evident that as we accumulate more definitive clinical information, ranging from history-taking to physical examination and imaging, there is an increase in entropy reduction useful for narrowing down differential diagnoses. To clarify, in our analysis, we operate under the assumption that when we calculate the removal of entropy, it is done in an independent manner, not simultaneously. This approach is due to our adoption of a decision tree model, which is structured to yield binary outcomes at each decision node. In other words, when encountering two or more findings or tests, we construct a decision tree. In this structure, the parent node initially holds the entropy, while the child node reflects the entropy after identifying the combination of findings. This approach enables us to calculate entropy removal effectively.

To effectively analyze two independent pieces of information in a sequential manner, we propose the construction of two interconnected decision trees. Initially, we assess the impact of elevated creatinine levels, calculating the entropy reduction (using TP, FP, FN, TN) associated with this condition alone. Subsequently, for the subset of individuals identified with high creatinine levels, we proceed to perform a CT scan to diagnose aortic dissection. At this second stage, we continue on calculating the True Positives (TP), False Positives (FP), False Negatives (FN), True Negatives

(TN) and subsequently the information gain for the presence of aortic dissection within this specific group.

Discussion

Extensive research has been conducted on the issue of diagnostic uncertainty in the medical literature (15). However, current methods aimed at addressing this problem, such as relying on clinical reasoning i.e. Evidence Based Medicine and utilizing metrics like sensitivity and specificity, lack flexibility and fail to comprehensively integrate patient-specific factors (16). These traditional approaches, although valuable, often fail to fully consider the rapidly changing and complex situations typically encountered in EM (17). Therefore, employing entropy and entropy removal, which quantify uncertainty, in specific scenarios such as aortic dissection in the emergency department, can underscore the advantages of utilizing them as metrics (6).

This is particularly valuable in situations where diagnostic uncertainty is high, such as in complex cases like aortic dissection. By considering the uncertainty inherent in medical diagnoses, entropy removal provides a more nuanced understanding of test performance and can better guide clinical decision-making. Additionally, entropy removal has the potential to adapt to changing patient conditions and evolving clinical scenarios, offering a more flexible and comprehensive approach to diagnostic assessment in the fast-paced and unpredictable environment of the ED. Thus, while sensitivity and specificity remain important metrics in medical diagnostics, entropy removal provides an additional complementary perspective that enhances our ability to navigate the complexities of emergency medicine and improve patient care.

In this study, we observed a positive correlation between entropy removal and both sensitivity and specificity, as detailed in Table 2. This result is expected from the equation of entropy removal which uses sensitivity and specificity as inputs. From an efficiency standpoint, entropy removal offers a single, unified metric that complements traditional measures such as sensitivity and specificity and can help synthesize their information into a single index of diagnostic uncertainty. This approach simplifies the determination of the most effective diagnostic tool by consolidating critical data into a single, comprehensive measure. Indeed, upon comparing the specificity and sensitivity of both MRI and Helical CT, it's evident that each modality significantly contributes to achieving an accurate diagnosis. However, when considering entropy removal as a metric, Helical CT demonstrates a greater capacity for entropy removal than MRI. This distinction suggests that, beyond conventional measures of diagnostic performance, Helical CT may offer enhanced effectiveness in reducing diagnostic uncertainty, thereby providing a more definitive analysis for certain medical conditions (18).

Moreover, entropy removal considers the sequence of diagnostic events, an aspect not captured by sensitivity and specificity. This feature is particularly valuable in emergency medicine, where the order of actions can critically influence patient outcomes. By accounting for the chronological progression of diagnostic interventions, entropy removal provides crucial insights into the most effective sequence of tests, thereby enhancing decision-making in high-stakes situations where time and accuracy are paramount. While PPV and NPV are intrinsically prevalence-dependent, we included them alongside sensitivity, specificity, and entropy removal to reflect how clinicians interpret test performance in real-world settings.

Nevertheless, there are some limitations to our study. Firstly, the dynamic nature of entropy, which can fluctuate based on the sequence of operations, remains unexplored in our analysis. This aspect holds significant potential for deepening our understanding of diagnostic accuracy and should be a focal point for subsequent investigations. Additionally, our analysis is based on

secondary diagnostic accuracy data derived from published studies and a curated diagnostic database. These studies differ in design, inclusion criteria, and clinical settings, which introduces heterogeneity into the underlying 2x2 tables. Such variability may influence both traditional metrics and entropy-based measures and limits the direct comparability of our results across tools and populations. Future studies should leverage harmonized datasets or electronic medical record-based cohorts to more rigorously evaluate entropy removal in real-world emergency department workflows.

By addressing these limitations, future research can enhance the precision and applicability of findings in the medical field.

References

- 1. Bartol, T. Thoughtful use of diagnostic testing: Making practical sense of sensitivity, specificity, and predictive value. *The Nurse Practitioner* **40**, 10-12 (2015).
 2. Naeger, D. M., Kohi, M. P., Webb, E. M., Phelps, A., Ordovas, K. G., & Newman, T. B..
- Correctly using sensitivity, specificity, and predictive values in clinical practice: how to avoid three common pitfalls. *American journal of roentgenology* **200**, W566-W570 (2013).
- 1. Juszczuk P, Kozak J, Dzikowski G, Głowania S, Jach T, Probierz B. Real-World Data Difficulty Estimation with the Use of Entropy. *Entropy* **23**, 1-36 (2021).
 2. Claude Shannon, 2021. "A Mathematical Theory of Communication (1948)", Ideas That Created the Future: Classic Papers of Computer Science, Harry R. Lewis.
 3. Harris KM, Nienaber CA, Peterson MD, et al. Early Mortality in Type A Acute Aortic Dissection: Insights From the International Registry of Acute Aortic Dissection. *JAMA Cardiol.* 2022;7(10):1009–1015. doi:10.1001/jamacardio.2022.2718
 4. He S, Chong P, Yoon BJ, Chung PH, Chen D, Marzouk S, Black KC, Sharp W, Safari P, Goldstein JN, Raja AS, Lee J. Entropy removal of medical diagnostics. *Sci Rep.* 2024 Jan 12;14(1):1181. doi: 10.1038/s41598-024-51268-4. PMID: 38216607; PMCID: PMC10786933.
- 1. Shiga T, Wajima Z, Apfel CC, Inoue T, Ohe Y. Diagnostic accuracy of transesophageal echocardiography, helical computed tomography, and magnetic resonance imaging for suspected thoracic aortic dissection: systematic review and meta-analysis. *Arch Intern Med.* 2006;166(13):1350-1356. doi:10.1001/archinte.166.13.1370
 2. Lee DK, Kim JH, Oh J, et al. Detection of acute thoracic aortic dissection based on plain chest radiography and a residual neural network (Resnet) [published correction appears in *Sci Rep.* 2023 Feb 9;13(1):2324]. *Sci Rep.* 2022;12(1):21884. Published 2022 Dec 19. doi:10.1038/s41598-022-26486-8
 3. von Kodolitsch Y, Nienaber CA, Dieckmann C, et al. Chest radiography for the diagnosis of acute aortic syndrome. *Am J Med.* 2004;116(2):73-77. doi:10.1016/j.amjmed.2003.08.030
 4. Nazerian P, Morello F, Vanni S, et al. Combined use of aortic dissection detection risk score and D-dimer in the diagnostic workup of suspected acute aortic dissection.

International journal of cardiology. 2014;175(1):78-82.

doi:10.1016/j.ijcard.2014.04.257

5. Asha SE, Miers JW. A Systematic Review and Meta-analysis of D-dimer as a Rule-out Test for Suspected Acute Aortic Dissection. *Annals of emergency medicine*. 2015;66(4):368-378. doi:10.1016/j.annemergmed.2015.02.013
6. Hata A, Yanagawa M, Yamagata K, et al. Deep learning algorithm for detection of aortic dissection on non-contrast-enhanced CT. *European radiology*. 2021;31(2):1151-1159. doi:10.1007/s00330-020-07213-w
7. von Kodolitsch Y, Schwartz AG, Nienaber CA. Clinical Prediction of Acute Aortic Dissection. *Arch Intern Med*. 2000;160(19):2977-2982. doi:10.1001/archinte.160.19.2977
8. Hammer MM, Kohlberg GD. Get the Diagnosis: an evidence-based medicine collaborative Wiki for diagnostic test accuracy. *Postgraduate Medical Journal* **93**, 179-185 (2017)
9. Bhise V, Rajan SS, Sittig DF, Morgan RO, Chaudhary P, Singh H. Defining and Measuring Diagnostic Uncertainty in Medicine: A Systematic Review. *J Gen Intern Med*. 2018 Jan;33(1):103-115. doi: 10.1007/s11606-017-4164-1. Epub 2017 Sep 21. PMID:

28936618; PMCID: PMC5756158.

•

1. de Alencar Neto, J.N., Santos-Neto, L. The Post Hoc Pitfall: Rethinking Sensitivity and Specificity in Clinical Practice. *J GEN INTERN MED* (2024). <https://doi.org/10.1007/s11606-024-08692-z>
2. Monaghan TF, Rahman SN, Agudelo CW, Wein AJ, Lazar JM, Everaert K, Dmochowski

RR. Foundational Statistical Principles in Medical Research: Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value. *Medicina (Kaunas)*. 2021 May 16;57(5):503. doi: 10.3390/medicina57050503. PMID: 34065637; PMCID: PMC8156826.

•

1. Sebastià C, Pallisa E, Quiroga S, Alvarez-Castells A, Dominguez R, Evangelista A. Aortic dissection: diagnosis and follow-up with helical CT. *Radiographics*. 1999;19(1):45-150. doi:10.1148/radiographics.19.1.g99ja0945