Artificial Intelligence in Healthcare: Systematic Review of Diagnostic and Screening Tools in Brazil's Resource-Limited Settings

Leticia Mancini, Luiz Eduardo Vanderlei Torres, Jorge Artur Peçanha de Miranda Coelho, Nichollas Botelho da Fonseca, Pedro Fellipe Dantas Cordeiro, Samara Silva Noronha Cavalcante, Diego Dermeval Medeiros da Cunha Matos

Submitted to: JMIR AI
on: December 04, 2024

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on it's website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressively prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Preprint
JMIR Publications

Artificial Intelligence in Healthcare: Systematic Review of Diagnostic and Screening Tools in Brazil's Resource-Limited Settings

Leticia Mancini[1]; Luiz Eduardo Vanderlei Torres[1]; Jorge Artur Peçanha de Miranda Coelho[1]; Nichollas Botelho da Fonseca[1]; Pedro Fellipe Dantas Cordeiro[1]; Samara Silva Noronha Cavalcante[1]; Diego Dermeval Medeiros da Cunha Matos[1]

[1]Universidade Federal de Alagoas Maceió BR

Corresponding Author:
Leticia Mancini
Universidade Federal de Alagoas
Av. Lourival Melo Mota, S/n - Tabuleiro do Martins
Maceió
BR

Abstract

Background: Artificial Intelligence (AI) has the potential to transform global healthcare, with extensive application in Brazil, particularly for diagnosis (D) and screening (S).

Objective: This study aimed to conduct a systematic review to understand AI applications in Brazilian healthcare, especially focusing on the resource-constrained environments.

Methods: A systematic review was performed. The search strategy included the following databases: PubMed, Cochrane Library, Embase, Web of Science, LILACS, and SciELO. The search covered articles from 1993 to November 2023, with 25 articles selected for the final sample. Meta-analysis data were evaluated based on three main metrics: ROC curve, sensitivity, and specificity. A random effects model was applied for each metric to address study variability.

Results: Key specialties for AI tools include ophthalmology and infectious disease, with a significant concentration of studies conducted in São Paulo state (52%). All articles included testing to evaluate and validate the tools; however, only two conducted secondary testing with a different population. In terms of risk of bias, 10 articles (40%) had medium risk, 8 articles (32%) had low risk, and 7 articles (28%) had high risk. Most studies were public initiatives, totaling 17 (68%), while 5 (20%) were private. In developing countries like Brazil, minimum technological requirements for implementing AI in healthcare must be carefully considered due to financial limitations and often insufficient technological infrastructure. Of the articles reviewed, 76% used computers, and 72% required the Windows operating system. The most used AI algorithm was Machine Learning (44%). The combined sensitivity was 0.8113, the combined specificity was 0.7417, and the combined area under the ROC curve (AUC) was 0.8308.

Conclusions: There is a relative balance in the use of both diagnostic and screening tools, with widespread application across Brazil in varied contexts. The need for secondary testing highlights opportunities for future research.

(JMIR Preprints 04/12/2024:69547)
DOI: https://doi.org/10.2196/preprints.69547

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✔ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.
2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?
✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Preprint
JMIR Publications

Original Manuscript

Preprint
JMIR Publications

Review

Artificial Intelligence in Healthcare: Systematic Review of Diagnostic and Screening Tools in Brazil's Resource-Limited Settings

Abstract

Background: Artificial Intelligence (AI) has the potential to transform global healthcare, with extensive application in Brazil, particularly for diagnosis (D) and screening (S).
Objective: This study aimed to conduct a systematic review to understand AI applications in Brazilian healthcare, especially focusing on the resource-constrained environments.
Methods: A systematic review was performed. The search strategy included the following databases: PubMed, Cochrane Library, Embase, Web of Science, LILACS, and SciELO. The search covered articles from 1993 to November 2023, with 25 articles selected for the final sample. Meta-analysis data were evaluated based on three main metrics: ROC curve, sensitivity, and specificity. A random effects model was applied for each metric to address study variability.
Results: Key specialties for AI tools include ophthalmology and infectious disease, with a significant concentration of studies conducted in São Paulo state (52%). All articles included testing to evaluate and validate the tools; however, only two conducted secondary testing with a different population. In terms of risk of bias, 10 articles (40%) had medium risk, 8 articles (32%) had low risk, and 7 articles (28%) had high risk. Most studies were public initiatives, totaling 17 (68%), while 5 (20%) were private. In developing countries like Brazil, minimum technological requirements for implementing AI in healthcare must be carefully considered due to financial limitations and often insufficient technological infrastructure. Of the articles reviewed, 76% used computers, and 72% required the Windows operating system. The most used AI algorithm was Machine Learning (44%). The combined sensitivity was 0.8113, the combined specificity was 0.7417, and the combined area under the ROC curve (AUC) was 0.8308.
Conclusions: There is a relative balance in the use of both diagnostic and screening tools, with widespread application across Brazil in varied contexts. The need for secondary testing highlights opportunities for future research.
Keywords: Artificial Intelligence; Brazil; Diagnosis; Screening.

Introduction

Artificial Intelligence (AI) can be defined in various ways, often simplified as an "imitation" of the human mind. However, these programs go beyond this definition, as they operate with different data sets and levels of autonomy, refined according to the goals set by developers. Given these technologies' vast applications, a significant transformation is occurring in the global health landscape. AI demonstrates utility in several areas, with some of its main competencies being differential D, lesion identification in imaging exams, and mortality prediction in hospitals [1-3].

In this context, it is crucial to recognize that the AI model developed is a significant element in the process, but the application environment is equally important. Using these tools in different environments depends on the resources available in each. Thus, it is necessary to consider how inequality in device access and the particularities of each location may interfere with the results, accuracy, and safety of the technologies employed [4].

Various models have been applied in Brazil, emphasizing diagnostic (D) and screening (S) areas. They are developed by the public and private sectors, despite 70% of the Brazilian population relying on the Unified Health System (SUS), which is the public healthcare system used in the country. It spans the entire territory and all levels of healthcare, ensuring comprehensive, universal, and free access to the entire population. Understanding how these technologies are used in sectors with more abundant resources can serve as a foundation for future implementations in SUS [5].

Therefore, this study aims to conduct a systematic literature review on the application of AI in health in Brazil, focusing on addressing the following questions: Is the tool used in the research for D or S? What is the context and location of the tool's application? Is the initiative public or private? Was the research funded? If so, by whom? What is the area/specialty of the tool? What type of AI application is used? What are the minimum requirements for using the tool? Was the tool tested? Was the tool tested on a population different from the one used to create the device? Was there evidence of health improvement?

These questions are fundamental for evaluating the effectiveness, accessibility, and safety of AI technologies in the Brazilian context as detailed in Table 1. Thus, the review allows us to explore and understand how these technologies are applied in Brazil and envision potential improvements for public health for a large part of this population.

Methods

Search Strategy and Selection Process

The search strategy was conducted in the following databases: PubMed, Cochrane Library, Embase, Web of Science, LILACS, and Scielo. Multimedia Appendix 1 presents the search strategies adapted for each database. The Rayyan application was used for duplicate study removal, and five different authors independently used it with the blind mode enabled to assess study selection [6]. The first phase involved analyzing the title, abstract, and keywords; at the end of this phase, if there was a disagreement, the authors met to discuss the application of inclusion and exclusion criteria for each article. The second phase involved a full-text reading. The searches identified articles from 1993 to November 2023, leaving 25 articles for the final sample.

Inclusion and Exclusion Criteria

The inclusion criteria for this study consisted of articles in which artificial intelligence is applied within the health sectors in Brazil as an aid for D and S of pathologies. Conversely, exclusion criteria were established to eliminate duplicate articles, literature reviews, animal studies, and studies where the application of artificial intelligence does not occur in the Brazilian context. Additionally, articles exclusively focused on robotic surgeries, AI use only within scientific methodology, hospital resource management, test reading, risk stratification and prognosis, risk factor assessment, epidemiological surveillance strategies, or any other topic not directly related to D and S of pathologies were excluded.

Data Extraction

Data extracted from each study included: authors and publication date, location and application context, purpose (D or S), type of initiative (public or private), study funding, benefiting specialty, type of AI application, whether tested in the population, tool functionality, and validation method. Additionally, values for accuracy, recall, precision, sensitivity, specificity, positive predictive value, and negative predictive value were obtained. Finally, the risk of bias was analyzed using the PROBAST (Prediction model Risk Of Bias Assessment Tool) (Table 3) to complete a risk assessment table based on the study data [7].

Statistical Analysis

Quantitative methods were adopted to perform the statistical analysis of the studies included in this systematic review, aiming to synthesize and interpret the collected data. First, a descriptive analysis was conducted on study characteristics, such as type of initiative (public or private), application areas, and type of AI used (D or S). These data were categorized and presented in absolute and relative frequencies.

Additionally, a quantitative analysis was conducted to assess the distribution of funding, classifying it as public, private, or mixed. The proportion of tested versus untested tools was evaluated to understand the robustness of the evidence presented by the studies. For issues related to application context and location, frequencies and percentages were calculated for different regions in Brazil. Finally, the analysis explored whether there was evidence of health improvement in studies that tested their tools.

Meta-analysis data were evaluated based on three main metrics: ROC Curve, Sensitivity, and Specificity. A random effects model was used for each metric to address the variability among the included studies. The exclusion of missing data was essential to ensure the quality and integrity of the results obtained.

Results

After searching the databases, 714 articles were found, of which 233 were from the PubMed database, 94 from EMBASE, 1 from Cochrane, 368 from Web of Science, six from Scielo, and 14 from Lilacs. After removing duplicates, 624 studies remained. In the first selection phase, applying the inclusion and exclusion criteria by analyzing the title, abstract, and keywords, 105 articles were selected for full-text reading. However, after removing inaccessible, incomplete articles or those not meeting the previously defined exclusion criteria, 84 articles remained. Subsequently, articles that did not involve the diagnostic and S process were removed, resulting in a final sample of 25 studies for this review, as illustrated in Figure 1. The articles are identified in table 2.

Table 1**:** Research Questions and Motivation

Research Questions	Description and Motivation
RQ1: Is the tool used for D or S?	This question helps clarify the primary objective of each intervention. This distinction is essential, as S focuses on maximizing sensitivity to ensure few actual cases are missed, while D aims to confirm or rule out a condition in individuals already identified as at risk, requiring a balance between sensitivity and specificity.
RQ2: What is the context and location of the tool's application?	The context of application is crucial to understanding the tool's target audience, as

	well as where and how it would be used. Additionally, the location indicates where these tools are more widely adopted and developed, highlighting the main centers advancing these technologies.
RQ3: Is the initiative public or private?	This question helps indicate whether the tool's purpose leans more towards improving the public health sector in the case of public initiatives or if private initiatives might focus more on cutting-edge technological innovation, operational efficiency, and potentially commercializing tools.
RQ4: Was the research funded? If so, by whom?	This question may influence the development, testing, implementation, and materials used for the tool.
RQ5: What is the field/specialty of the tool?	The tool's field is essential to understand the specific needs of that specialty, ranging from data requirements to expected outcomes.
RQ6: What type of AI application is it?	This question clarifies the technological approach adopted and its implications for the tool's performance. Different types of AI have distinct capabilities and limitations that can directly impact the tool's accuracy, adaptability, and complexity. This choice affects model robustness, data types analyzed, and, consequently, tool effectiveness.
RQ7: What are the minimum requirements for using the tool?	This question is essential for assessing implementation feasibility in developing contexts like Brazil, as financial and technological constraints are crucial to ensuring effective and accessible tool usage.

RQ8: Was the tool tested?	This question is critical, as tool testing is necessary to validate its efficacy, accuracy, and safety before clinical implementation.
RQ9: Was the tool tested on a population different from the one used to create the device?	This question assesses the tool's applicability across different contexts and populations, helping avoid sample bias and ensure functionality across diverse settings.

Table 2: Identification

Article Identification	Article Title
1	DZC DIAG: mobile application based on expert system to aid in the diagnosis of dengue, Zika, and chikungunya
2	Prediction of metabolic syndrome: A machine learning approach to help primary prevention
3	Impact of radiomics on the breast ultrasound radiologist’s clinical practice: from lumpologist to data wrangler
4	Automated algorithms combining structure and function outperform general ophthalmologists in diagnosing gl oma
5	Computer-aided diagnosis system based on fuzzy logic for breast cancer categorization
6	Artificial Intelligence in Allergy and Immunology: Comparing Risk Prediction Models to Help Screen Inborn Errors of Immunity.
7	Artificial intelligence techniques applied to the development of a decision-support system for diagnosing celiac disease
8	An online platform for COVID-19 diagnostic screening using a machine learning algorithm
9	Clinical validation of a smartphone-based retinal camera for diabetic retinopathy screening
10	Covid-19 Automated Diagnosis and Risk Assessment through Metabolomics and Machine Learning
11	NICeSim: an open-source simulator based on machine learning techniques to support medical research on prenatal and perinatal care decision making
12	Screening for Chagas disease from the electrocardiogram using a deep neural network
13	Left ventricular systolic dysfunction predicted by artificial intelligence using the electrocardiogram in Chagas disease patients–The SaMi-Trop cohort.
14	Implementation of an expert system to determine eligibility and priorities for bone marrow transplants
15	Multi-center Integrating Radiomics, Structured Reports, and Machine Learning Algorithms for Assisted Classification of COVID-19 in Lung Computed Tomography

16	Screening for active pulmonary tuberculosis: Development and applicability of artificial neural network models
17	Can machine learning be useful as a screening tool for depression in primary care?
18	Diabetic Retinopathy Screening Using Artificial Intelligence and Handheld Smartphone-BasedRetinal Camera
19	Implementation of artificial intelligence algorithms for melanoma Screening in a primary care setting
20	Decision support system for the diagnosis of schizophrenia disorders
21	Leprosy Screening Based on Artificial Intelligence
22	Single retinal image for diabetic retinopathy
24	Artificial neural networks applied to study allergic conjunctivitis
25	Osteoporosis screening using machine learning and electromagnetic waves
26	A screening system for smear-negative pulmonary tuberculosis

Risk of Bias

The assessment of risk of bias in the articles revealed the following classifications (Figure 2): 7 articles (28.00%) presented medium risk, 5 articles (20.00%) presented low risk, and 13 articles (52.00%) presented high risk [8-32]. Among the articles with high risk, 6 (46.15%) showed high risk in data analysis [24, 25, 26, 29, 31, 32]. Article 11 stood out with a high risk of bias in analyses, predictors, and outcomes [25].

Table 3: (Risk of Bias Questions)

	Question	Possible answers
QV1	Were appropriate data sources used, for example, cohort, randomised controlled trial, or nested case-control study data?	Y N ?
QV2	Were all inclusions and exclusions of participants appropriate?	Y N ?
QV3	Were predictors defined and assessed in a similar way for all participants?	Y N ?
QV4	Were predictor assessments made without knowledge of outcome data?	Y N ?
QV5	Were all predictors available at the time the model was intended to be used?	Y N ?
QV6	Was the outcome determined appropriately?	Y N ?
QV7	Was a prespecified or standard outcome definition used?	Y N ?
QV8	Were predictors excluded from the outcome definition?	Y N ?
QV9	Was the outcome defined and determined in a similar way for all participants?	Y N ?
QV10	Was the outcome determined without knowledge of predictor information?	Y N ?

QV11	Was the time interval between predictor assessment and outcome determination?	Y N ?
QV12	Were there a reasonable number of participants with the outcome?	Y N ?
QV13	Were continuous and categorical predictors handled appropriately?	Y N ?
QV14	Were all enrolled participants included in the analysis?	Y N ?
QV15	Were participants with missing data handled appropriately?	Y N ?
QV16	Was selection of predictors based on univariable analysis avoided?	Y N ?
QV17	Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately?	Y N ?
QV18	Were relevant model performance measures evaluated appropriately?	Y N ?
QV19	Were model overfitting and optimism in model performance accounted for?	Y N ?

Y = Yes, N = No, and ? = Without information

Participants

The lowest risk of bias among articles was in participant selection. For participant selection, 14 (56.00%) presented low risk (2,3,4,6,12,13,14,15,16,17,20,23,24,25), 8 (32.00%) presented medium risk (1,9,10,11,18,19,21,22), and 3 (12.00%) presented high risk (5,7,8) [8-32].

Outcomes

For outcomes, 13 (52.00%) presented low risk (2,4,5,6,7,8,9,16,17,18,19,21,22), 9 (36.00%) presented medium risk (1,3,12,13,14,15,20,24,25), and 3 (12.00%) presented high risk (10, 11, and 23) [8-32].

Analysis

The criterion with the highest risk of bias was data analysis, with 13 (52.00%) articles showing high risk (3,5,6,8,10,11,14,15,20,21,22,23,25), 7 (28.00%) presenting medium risk (1,4,7,13,16,25), and 5 (20.00%) presenting low risk (2,9,12,17,19) [8-32].

Predictors

For predictors, 7 (28.00%) presented low risk (4,5,8,9,10,15,17), 12 (48.00%) presented medium risk (1,2,3,6,12,13,16,18,19,20,22,23), and 6 (24.00%) presented high risk (7,11,14,21,24,25) [8-32].

Article Objective

The articles' objectives were classified as D or S. Most (13; 52.00%) were focused on D, while 12 (48.00%) on S [8-32].

Specialty

The medical areas covered by the tools are varied, with ophthalmology predominating, being the subject of 5 (20.00%) articles, with an emphasis on diabetic retinopathy D [9, 13, 14, 30, 31]. Next is infectology, covered in 4 (16.00%) articles, focusing on COVID-19 D [8, 23, 24, 27]. Additionally, specialties such as cardiology, internal medicine, dermatology, pulmonology, and mastology are each the subject of 2 (8.00%) articles, totaling 40.00% of the articles [11, 12, 14, 17, 18, 19, 20, 21, 29, 32]. Other specialties, such as allergy and immunology, endocrinology and metabolism, gastroenterology, neonatology, oncology, and psychiatry, are each covered in 1 (4.00%) article, totaling 24.00% of the articles [10, 15, 22, 25, 26, 28],.

Public/Private Initiative and Funding

The selected articles were characterized according to the nature of their initiative, being public or private. Most were part of a public initiative, totaling 17 (42.00%), while 5 (20.00%) were private initiatives, and 1 (4.00%) did not provide this information. Regarding funding, 7 (28.00%) articles did not provide information on this issue. Two (8.00%) articles were funded solely by the São Paulo Research Foundation (FAPESP), 2 (8.00%) by FAPESP in association with the National Council for Scientific and Technological Development (CNPQ), 1 (4.00%) by FAPESP in conjunction with various laboratories, 1 (4.00%) by FAPESP with the Technological Development Support Laboratory (LADETEC), 3 (12.00%) by the Minas Gerais State Research Support Foundation (FAPEMIG) and CNPQ, and 1 (4.00%) by both and CAPES. Additionally, 1 (4.00%) was funded by CNPQ and the National Institutes of Health (NIH), 1 (4.00%) by MIT libraries, 1 (4.00%) by the Pontifical Catholic University of Paraná (PUCPR), 1 (4.00%) by FINep and CAPES, and 2 (8.00%) by FAPERJ, with one (50.00%) in conjunction with ICNT and CNPQ and the other (50.00%) with CAPES. Finally, 1 (4.00%) was funded by PROADI-SUS, and 1 (4.00%) had no financial support. This information demonstrates a predominance of funding, respectively, from CNPQ and FAPESP [8-32].

Application Context/Location

Almost half of the articles were conducted in the state of São Paulo, totaling 12 (48.00%), of which 3 (25.00%) were conducted in collaboration with the states of Sergipe, Minas Gerais, and Amazonas [9, 10, 11, 15, 16, 19, 20, 21, 22, 24, 28, 31]. Among these, 4 (33.33%) were associated with the Hospital das Clínicas of the University of São Paulo [10, 21, 24, 28]. One of the studies also included the Hospital Estadual Sumaré, the Municipal Hospital of Paulínia, and the Delphina Rinaldi Abdel Aziz Hospital in Manaus [24]. The hospitals associated with Sergipe were unspecified, and the study in Minas Gerais was conducted at the Tropical Medicine Research Center in both states [11, 16].

The other 4 (33.33%) studies conducted in the state of São Paulo took place at the São Camilo University Center, the Albert Einstein Israeli Hospital, the UNIFESP outpatient clinic, and an unspecified private healthcare institution [15, 19, 20, 31].

Two (8.00%) tools from the articles were applied to populations during campaigns about Diabetes Mellitus, one in Santa Catarina and the other in Bahia [13, 30]. Three (12.00%) studies were conducted in Rio de Janeiro, one (33,33%) at the UFRJ University Hospital , other (33,33%) at the Augusto Amaral Peixoto Polyclinic and the other at Santa Casa de São Sebastião (33,33%) [12, 25, 32]. Three (12.00%) studies were conducted at the University Hospitals of UFRN (Rio Grande do Norte), UFPR (Paraná), and UFPE (Pernambuco) [8, 14, 26]. One (4.00%) study was conducted at the University Hospitals of Bahia, Paraíba, and Minas Gerais, associated with their respective federal universities [27]. Another 4 (16.00%) studies used data from the internet, with one (25.00%) associated with UFMG, two (50.00%) with UFRJ, and one (25.00%) with UFJF [17, 18, 23, 29].

Artificial Intelligence Algorithms

The analyzed articles used different Artificial Intelligence algorithms, classified into five distinct groups, with some studies combining more than one type in their development (Table 4). Eleven (44.00%) studies used Machine Learning, nine (36.00%) employed Deep Learning, one applied Fuzzy Logic, three (12.00%) adopted Ensemble Methods, and three (12.00%) utilized Expert

Systems [8-32].

Table 4: Tipe of IA and Minimum Requirements [a]

Article Identification	Type of IA	Minimum Requirements
1	ES	Cell phone with sufficient memory and Android software version 4.0 (Ice Cream Sandwich) or higher.
2	ML, EM	Computer compatible with Windows, with sufficient memory and compatibility to run the Machine Learning models used.
3	ML	Computer compatible with Windows, with sufficient memory and compatibility to run the Machine Learning models used.
4	DL, ML, EM	Computer compatible with Windows, with sufficient memory and compatibility to run the Machine Learning models used.
5	FL	Computer compatible with Windows, with sufficient memory and compatibility to run the Machine Learning models used.
6	ML	Computer compatible with Windows, with sufficient memory and compatibility to run the Machine Learning models used.
7	ML	Computer compatible with Windows, with sufficient memory and compatibility to run the Machine Learning models used.
8	ML	Computer compatible with Windows, with sufficient memory and compatibility to run the Machine Learning models used.
9	DL	Specific cell phone model: Samsung Galaxy S10 (Android 11)
10	ML	Computer compatible with Windows, with sufficient memory and compatibility to run the Machine Learning models used.
11	ML	Computer with specific software (NICeSim).
12	DL	Windows-compatible computer.
13	DL	ECG device connected to a Windows-compatible computer.
14	ES	Windows-compatible computer (Expert Sinta Software).
15	EM	Windows-compatible computer.
16	ML	Computer compatible with Windows, with sufficient memory and compatibility to run the Machine Learning models used.
17	ML	Computer compatible with Windows, with sufficient memory and compatibility to run the Machine Learning models used.
18	DL	Specific cell phone model: Samsung Galaxy S10 (Android 11)
19	DL	Android and iOS platforms for applicability, with no specific model requirement.
20	ES	Not specified
21	ML	Android and iOS platforms for applicability, with no specific model requirement.
22	DL	Computer compatible with Windows, with sufficient memory and compatibility to run the Machine Learning models used.
24	DL	Computer compatible with Windows, with sufficient memory and compatibility to run the Machine Learning models used.
25	ML	Computer compatible with Windows, with sufficient memory and compatibility to run the Machine Learning models used.
26	DL	Computer compatible with Windows, with sufficient memory and compatibility to run the Machine Learning models used.

[a]Expert Systems = ES; Machine Learning = ML; Ensemble Methods = EM; Deep learning = DL; Fuzzy Logic = FL.

Minimum Requirements

Given the context of developing countries, such as Brazil, and their financial limitations, the minimum technological requirements necessary for the functionality of the tools were observed. The primary requirement is a computer, required by 19 (76.00%) of the articles, with 16 (84.21%) needing memory and compatibility for the Machine Learning models in these studies, 2 (10.53%) requiring specific software, and 1 (5.26%) with an ECG device connected to the computer, all compatible with Windows [9, 10, 11, 12, 14 , 15, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 30, 31, 32].

Additionally, 3 (12.00%) required specific cell phone models: a phone with sufficient memory and Android software version 4.0 (Ice Cream Sandwich) (article 1) and Samsung Galaxy S10 (Android 11) (articles 9 and 18) [8, 13, 14]. Two (8.00%) were applicable on Android and IOS platforms, with no specific device model requirements (articles 19 and 21)[19, 29]. One (4.00%) article did not specify the necessary technology [28].

Test / External Application / Evidence of Health Improvement

All the articles conducted some form of testing to evaluate and validate the tools. However, only two of them performed a second evaluation with application to a different population. Additionally, all studies presented some evidence of health improvement, but the risks of bias associated with these evidences will be explored [8-32].

Results of Statistical Analyses

The sensitivity meta-analysis, using a random-effects model, resulted in a combined estimate of 0.8113 (95% CI: 0.7856 - 0.8369), with a Z value of 62.0035 ($p < 0.0001$). The analysis revealed high heterogeneity among the studies ($I^2 = 88.23\%$), suggesting significant variability in the results of the included studies. The forest plot in Figure 3 visualizes each study's sensitivity estimate and the combined estimate, highlighting the dispersion of the results.

The combined specificity estimate was 0.7417 (95% CI: 0.7113 - 0.7722), with a Z value of 47.7605 ($p < 0.0001$). The analysis indicated high heterogeneity ($I^2 = 88.92\%$), as shown in Figure 4, reflecting substantial differences in the study results. The corresponding forest plot illustrates these individual estimates and the combined estimate from the meta-analysis.

The ROC Curve meta-analysis presented a combined estimate of 0.8308 (95% CI: 0.8124 - 0.8492), with a Z value of 88.4640 ($p < 0.0001$). The heterogeneity was moderate to high ($I^2 = 76.52\%$), as visualized in the forest plot in Figure 5, suggesting variations in methods or populations among the studies.

The forest plots (Figures 3, 4 and 5) provide a visual representation of the individual and combined estimates for sensitivity, specificity, and ROC Curve. Each graph highlights the variability among the studies and the combined estimate obtained through the meta-analysis. The dashed vertical line in the graphs represents the combined estimate, while the confidence intervals are indicated by horizontal bars for each study.

Discussion

Principal Results

The application of AI in healthcare has advanced significantly in Brazil, with various tools being developed for S and D purposes. This distinction between S and D is essential, as it directly influences the implementation and impact of technologies in different healthcare contexts. The

Preprint JMIR Publications

primary goal of S tools is to identify individuals at risk of a given condition, prioritizing sensitivity maximization. This means these tools are designed to ensure that few actual cases go unnoticed, even if it increases false positives. This approach is fundamental in scenarios where early detection can save lives or prevent disease progression [33].

On the other hand, D tools aim to confirm or rule out the presence of a condition in individuals already identified as at risk. In these cases, accuracy becomes crucial, with a more careful balance between sensitivity and specificity. Often, there is a focus on maximizing specificity to reduce false positives, thus avoiding unnecessary treatments. These accuracy requirements reflect the distinct needs of S and D tools, with S being more permissive regarding false positives to ensure no actual case is overlooked, while D seeks high accuracy to prevent clinical errors [33].

The implementation of these tools also varies depending on the context of use. S tools are often used in resource-limited settings, where there is a need to process large volumes of data quickly. Thus, they are utilized in community S campaigns or public health programs, where the focus is on quickly identifying cases that need more thorough evaluation, as shown in articles: 6, 9, 16, 18, 22, and 26 [12, 13, 16, 22, 23, 32] . Diabetic retinopathy S is prominent in this context, covered in articles 9, 18, and 22, as well as tuberculosis in articles 16 and 26, conditions highlighted in the Brazilian context [12, 13, 16, 30, 32]. Rare syndromes that require early intervention, such as inborn immunity errors in article 6, were also explored [22].

Highlighted specialties for AI tools include ophthalmology and infectious disease. The former benefits from AI’s image analysis capabilities, revisiting the topic discussed earlier: diabetic retinopathy research (9, 18, 22), allergic conjunctivitis S (24), and glaucoma D (4) [9, 13, 16, 30, 31]. Meanwhile, infectious disease, critically important in Brazil, stood out in COVID D—a crucial aspect, given the global impact of the pandemic. In this case, tools were used for clinical D (8, 10), with one utilizing imaging diagnostics through computed tomography (15) [23, 24, 27]. Another area worth noting, not due to the number of tools produced but because of its clinical importance in Brazil, also within infectious disease, is pulmonology, which includes tuberculosis S, with both articles covering testing and clinical symptoms (16 and 26) [12, 32].

The predominance of public initiatives and funding from scientific development programs—particularly National Council for Scientific and Technological Development (CNPQ) and São Paulo Research Foundation (FAPESP)—highlights the importance of state support for research. This support is crucial for advancing tools and studies in healthcare, significantly benefiting the country. Furthermore, the presence of private initiatives reflects the ongoing need for healthcare technology advancement in Brazil. Public-private partnerships, exemplified by PROADI-SUS, are essential for improving practices and innovations in the sector, fostering a constantly evolving environment [8-32].

Brazil’s vast and diverse territory shows varying trends in areas such as infrastructure, access to technology, and development in healthcare and educational institutions [34]. This review revealed a significant concentration of studies conducted in São Paulo state (48%), with the Hospital das Clínicas at the University of São Paulo being involved in 33.33% of these studies [9, 10, 11, 15, 16, 19, 20, 21, 22, 24, 28, 31]. This event highlights the tendency for AI usage to be concentrated in the country's wealthiest state: São Paulo. Additionally, a notable collaboration between São Paulo and other states, such as Sergipe, Minas Gerais, and Amazonas, highlights joint efforts to develop technology across regional boundaries within the country [11, 16, 24].

This is essential as it allows adaptation to different realities, with broader applications and more

effective tools (tested in various populations).
Other states, such as Rio Grande do Norte, Bahia, Santa Catarina, and Pernambuco,
also produced AI tools, demonstrating broad participation across Brazilian states [8, 13, 14, 30]. Finally, prominent institutions, including Federal University of Minas Gerais (UFMG), Federal University of Rio de Janeiro (UFRJ), and Federal University of Juiz de Fora (UFJF), used internet data to develop their technologies, showcasing the importance of digital resources for health research, expanding the reach and diversity of studied populations [17, 18, 23, 29].

The most commonly used AI algorithm was Machine Learning (44%), which employs models attempting to predict future outcomes based on a dataset and is helpful in automated diagnostics and decision-support systems, as seen in articles 7 and 25 [10, 14, 35]. Deep Learning, a subcategory of Machine Learning, appeared in 36% of the studies, allowing more complex data analysis involving larger volumes and image recognition, a critical feature in articles 18 and 22, for example, for image evaluation [13, 23, 36] . Additionally, article 15 used Ensemble Methods, which combine multiple machine learning models to improve decision accuracy [27, 37].

Another important algorithm is Expert Systems, which operate based on a predefined set of rules and knowledge, used for decision support, as in article 14, or D, as in article 1 [8, 26]. Finally, Fuzzy Logic is valuable for decision-making in less precise cases requiring flexible interpretation, such as breast cancer characterization in Article 5 [21].

In developing countries like Brazil, minimum technological requirements for implementing AI tools in healthcare should be carefully assessed due to financial constraints and frequently inadequate technological infrastructure. The analysis of minimum requirements, as presented, reveals a predominant reliance on computers as a fundamental platform, with 76% of articles indicating their necessity, underscoring the importance of computing in executing Machine Learning models [9, 10, 11, 12, 14, 15, 17, 18, 20, 21, 22 , 23, 24, 25, 26, 27, 30, 31, 32].

The need for computers compatible with Windows operating systems was observed in 72.21% of the articles, revealing this operating system's prevalence in healthcare environments in Brazil. The use of specific software for running AI models was also noted, though at a lower rate (10.53%). The compatibility with Windows may be attributed to its wide adoption and familiarity in the Brazilian market, potentially reducing barriers to adopting new technologies [9, 10, 11, 14, 17, 20, 22 , 23, 24, 25, 26, 27, 30, 31, 32].

Besides computers, one article mentioned using ECG devices linked to computers, indicating additional hardware integration for data monitoring and collection [14]. The need for specific mobile devices was identified in 12% of the articles, focusing on phones with minimum specifications, such as sufficient memory and specific Android OS versions. This requirement for compatibility with particular versions may limit the applicability of tools on older or less common mobile devices, potentially restricting technology access and dissemination, a negative point.
The diverse application of tools across Android and iOS platforms, without strict device model specifications, reflects an attempt at greater flexibility and accessibility [19, 29].

The meta-analysis on AI model performance in medical diagnostics and S revealed important clinical insights. The combined sensitivity was 0.8113, indicating that, on average, AI models correctly identify approximately 81% of true positives. This high sensitivity is crucial in clinical settings, particularly in initial S, as it helps ensure that most patients with a medical condition are identified and treated, minimizing the risk of false negatives.

The combined specificity was 0.7417, suggesting that AI models correctly identify around 74% of true negatives. Although this specificity is considered good, it is lower than sensitivity, which may indicate a tendency for false positives in some scenarios. This is relevant, as false positives can lead to unnecessary treatments or additional tests for patients who do not have the condition in question. Due to existing financial constraints, these potential unnecessary costs are harmful, especially in low-resource settings.

The combined area under the ROC curve (AUC) was 0.8308, pointing to AI models' excellent discriminatory ability between positive and negative cases. The AUC reflects overall model performance, demonstrating a satisfactory balance between sensitivity and specificity, which is essential for accurate diagnostics and S.

Limitations

All articles presented evidence of health improvement, suggesting that AI tools have the potential to positively impact clinical outcomes. However, testing is crucial to ensure the effectiveness and safety of AI tools in healthcare. All articles reviewed reported some form of testing to validate their tools, indicating a commitment to performance evaluation of the proposed solutions.

Nevertheless, a major limitation of the study is that only two articles performed a second evaluation by applying the tool to a different population, limiting the generalizability of the results and increasing the possibility of statistical errors in real-world population outcomes.

Another relevant factor is the high heterogeneity observed in all analyses, with I^2 values ranging from 76.52% to 88.92%. This variability suggests that the effectiveness of AI models may vary significantly between different studies and contexts. Differences in algorithms used, study populations, and implementation methods may contribute to this variation, further reinforcing the concern about retesting these tools in other contexts.

Additionally, the risk of bias presented, with many articles showing moderate to high risk, highlights the need for a more transparent methodology in these studies to evaluate the tools more safely and effectively, reflecting real-world applications.

Conclusions

This review demonstrates a broad application of AI technologies in diagnostic and treatment areas, with relative equality in the use of both types of tools and broad usage across Brazil in varied contexts. The predominance of public funding indicates the potential for tools for use in the Brazilian SUS, which is significant due to the widespread reliance on the Brazilian public healthcare system. The variety of specialties highlights the diversity of AI applications and their importance in the health sector. Finally, the need for secondary testing points to future research opportunities.

Acknowledgements

This research was funded by the National Council for Scientific and Technological Development (CNPq). The author(s) gratefully acknowledge the support provided, which made this study possible.

Conflicts of Interest

The authors have no competing interests to declare that are relevant.

Preprint
JMIR Publications

Abbreviations

AI: Artificial Intelligence
D: Diagnosis
S: Screening
SUS: Unified Health System

References

1. Ciecierski-Holmes T, Tretiakov A, Lam D, Tanaka MJ, Goldenberg A, Nair R, et al. Artificial intelligence for strengthening healthcare systems in low- and middle-income countries: a systematic scoping review. NPJ Digit Med; 2022;5(1):162. doi:10.1038/s41746-022-00700-y

2. World Health Organization. Ethics and governance of artificial intelligence for health. Available from: https://www.who.int/publications/i/item/9789240029200. ISBN:9789240029200.

3. Saif-ur-Rahman KM, Basnet B, Chakravarthy S, van Loenhout JAF, Müller L, Elhadi M, et al. Artificial intelligence and digital health in improving primary health care service delivery in LMICs: a systematic review. J Evid Based Med; 2023;16(3):303-20. doi:10.1111/jebm.12529

4. Alami H, Rivard L, Lehoux P, Hoffman SJ, Cadeddu SB, Savoldelli M, et al. Artificial intelligence in health care: laying the foundation for responsible, sustainable, and inclusive innovation in low- and middle-income countries. Glob Health;2020;16(1):52. doi:10.1186/s12992-020-00584-1

5. Pesquisa Nacional de Saúde 2019. Informações sobre domicílios, acesso e utilização dos serviços de saúde. Brasil, Grandes Regiões e Unidades da Federação [Internet]. Rio de Janeiro: IBGE; 2019. Available from: https://biblioteca.ibge.gov.br/visualizacao/livros/liv101748.pdf.28.

6. Rayyan Systems Inc. Rayyan: A systematic reviews web app. Available at: https://www.rayyan.ai/

7. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, Reitsma JB, Kleijnen J, Mallett S; PROBAST Group. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med. 2019;170(1):51-58. doi: 10.7326/M18-1376. PMID: 30596875.

8. de Araújo APR, Bezerra CG, Santos LAC, Mota AS, Lima FM, Furtado A, et al. DZC DIAG: mobile application based on expert system to aid in the diagnosis of dengue, Zika, and chikungunya. Med Biol Eng Comput; 2020;58(11):2657-72. doi:10.1007/s11517-020-02233-6

9. Shigueoka LS, Fujino AY, Lopez PF, Fernandes YB, Gomes GF, Passos CJ, et al. Automated algorithms combining structure and function outperform general ophthalmologists in diagnosing glaucoma. PLoS One; 2018;13(12):0207784. doi:10.1371/journal.pone.0207784

10. Tenório JM, Reis AD, Lopes FS, Santos RP, Diniz PR. Artificial intelligence techniques applied to the development of a decision-support system for diagnosing celiac disease. Int J Med Inform. 2011;80(11):793-802. doi:10.1016/j.ijmedinf.2011.09.003

11. Oliveira B, Souza FS, Ribeiro TL, Almeida GC, Barros N. Left ventricular systolic dysfunction predicted by artificial intelligence using the electrocardiogram in Chagas disease patients–The SaMi-Trop cohort. PLoS Negl Trop Dis. 2021 ;15(12):0009974. doi:10.1371/journal.pntd.0009974

12. Souza Filho JBO, Cardoso JS, Lima RC, Santana GP, Cássio LA, Silva J. Screening for active pulmonary tuberculosis: development and applicability of artificial neural network models. Tuberculosis (Edinb); 2018;111:94-101. doi:10.1016/j.tube.2018.06.004

13. Malerbi FK, Lima VL, Mesquita MR, Tavares W, Rocha HA, Maciel PA, et al. Diabetic retinopathy screening using artificial intelligence and handheld smartphone-based retinal camera. J Diabetes Sci Technol; 2021;16(3):716-23. doi:10.1177/19322968211024402

14. Albuquerque GA, de Sousa RG, Barbosa BT, Lima RCG, Carneiro JF, Soares FJ. Osteoporosis screening using machine learning and electromagnetic waves. Sci Rep; 2023;13(1):12865. doi:10.1038/s41598-023-40104-w

15. Tavares LD, Garcia RC, Alves M, Nascimento R, Cangussu D, Tavares W. Prediction of metabolic syndrome: a machine learning approach to help primary prevention. Diabetes Res Clin Pract; 2022;191:110047. doi:10.1016/j.diabres.2022.110047

16. Angélica J, Lima ML, Garcia M, Macêdo MM, Santana GA. Clinical validation of a smartphone-based retinal camera for diabetic retinopathy screening. Acta Diabetol; 2023;60(8):1075-81. doi:10.1007/s00592-023-02105-z

17. Jidling C, Nygren L, Subrahmanian P, Hamza EA, Sajid MI, Dayoub R, et al. Screening for Chagas disease from the electrocardiogram using a deep neural network. PLoS Negl Trop Dis; 2023;17(7):0011118. doi:10.1371/journal.pntd.0011118

18. Souza Filho EM, Nakazato C, Gonçalves PC, Amado D, Silva AG. Can machine learning be useful as a screening tool for depression in primary care? J Psychiatr Res; 2021;132:1-6. doi:10.1016/j.jpsychires.2020.10.017

19. Giavina-Bianchi M, Azevedo-Marques P, Veronesi F, Otta TM, Vugman I, Arias S, et al. Implementation of artificial intelligence algorithms for melanoma screening in a primary care setting. PLoS One; 2021;16(9):e0257006. doi: 10.1371/journal.pone.0257006

20. Fleury EFC, Marcomini K. Impact of radiomics on the breast ultrasound radiologist's clinical practice: from lumpologist to data wrangler. Eur J Radiol; 2020 ;131:109197. doi:10.1016/j.ejrad.2020.109197

21. Miranda GHB, Felipe JC. Computer-aided diagnosis system based on fuzzy logic for breast cancer categorization. Comput Biol Med; 2015 ;64:334-46. doi:10.1016/j.compbiomed.2015.02.015

22. Takao MMV, Silva JB, Zucchi P. Artificial intelligence in allergy and immunology: comparing risk prediction models to help screen inborn errors of immunity. Int Arch Allergy Immunol; 2022 ;1-5. doi:10.1159/000525339

23. Marques E, da Silva RG, Pereira CC. An online platform for COVID-19 diagnostic screening using a machine learning algorithm. J Infect Dis; 2023;69(4):13468.

24. Delafiori J, Navarro LC, Val FF, Garcia JS, Moritz RL. COVID-19 automated diagnosis and risk assessment through metabolomics and machine learning. Anal Chem; 2021;93(4):2471-9. doi:10.1021/acs.analchem.0c04497

25. Cerqueira FR, Reis LP, Sousa LP, Rolo RC. NICeSim: an open-source simulator based on machine learning techniques to support medical research on prenatal and perinatal care decision making. Med Biol Eng Comput; 2014;62(3):193-201. doi: 10.1016/j.artmed.2014.10.001

26. Dias S. Implementation of an expert system to determine eligibility and priorities for bone marrow transplants. Stud Health Technol Inform; 2020;84(Pt 1):293-6. PMID: 11604773

27. Marcos LG, Lourenço RT, Pereira VJ, Meireles RA, Vasconcelos MC, Pitta AM, et al. Multi-center integrating radiomics, structured reports, and machine learning algorithms for assisted classification of COVID-19 in lung computed tomography. J Med Biol Eng; 2023;43(2):156-62. doi:10.1007/s40846-023-00781-4

28. Razzouk D, Zullo K, Marciano AE, Dos Santos AB, Caixeta L, Maeda MY, et al. Decision support system for the diagnosis of schizophrenia disorders. Braz J Med Biol Res; 2006;39(1):119-28. doi:10.1590/S0100-879X2006000100016

29. de Souza MLM, Silva MF, Gonçalves F, de Souza N, Yassuda P. Leprosy screening based on artificial intelligence: development of a cross-platform app. JMIR Mhealth Uhealth; 2021;9(4):e23718. doi:10.2196/23718

30. Penha FM, Gawryszewski LG, Bittencourt BB, Tavares JM. Single retinal image for diabetic retinopathy screening: performance of a handheld device with embedded artificial intelligence. Ophthalmol Retina; 2023;9(1):e12865. doi:10.1186/s40942-023-00477-6

31. Goulart DA, Souza RT, Garcia FG, Silva TP, Sant'Anna LP. Redes neurais artificiais aplicadas no estudo de questionário de varredura para conjuntivite alérgica em escolares. Arq Bras Oftalmol; 2006;69(5):707-13.

32. de O Souza Filho JB, Scherr B, Carniel E, Pessanha A, Ribeiro Júnior B, Marchioro D, et al. A screening system for smear-negative pulmonary tuberculosis using artificial neural networks. Int J Infect Dis; 2016;49:33-9. doi: 10.1016/j.ijid.2016.05.019

33. Warner J. Clinicians' guide to evaluating diagnostic and screening tests in psychiatry. Psychiatry;2000;62(2-3):165-173. doi:10.1192/apt.10.6.446

34. Ibanez P, Westmann G, Iozzi FL. Territorial planning and regionalization in Brazil: empirical consolidation and the role of the national development bank. Springer; 2021;113-26.

35. Domingos P. A few useful things to know about machine learning. Commun ACM; 2012;55(10):78. doi:10.1145/2347736.2347755

36. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature; 2015 ;521(7553):436-44. doi:10.1038/nature14539

37. Dietterich TG. Ensemble methods in machine learning. Multiple Classifier Systems; 2000;1857:1-15. doi:10.1007/3-540-45014-9_1

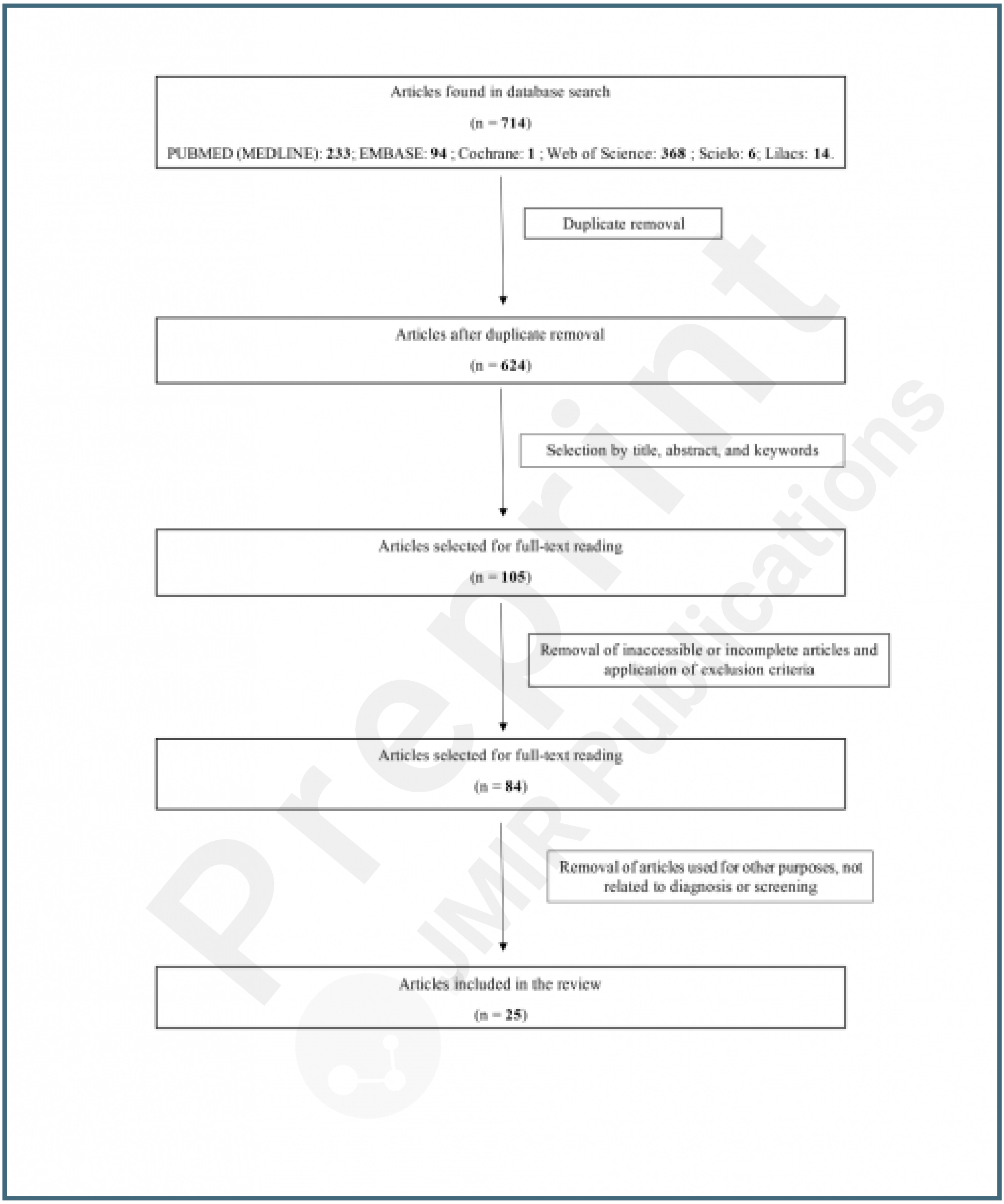
Preprint JMIR Publications

Supplementary Files

Preprint
JMIR Publications

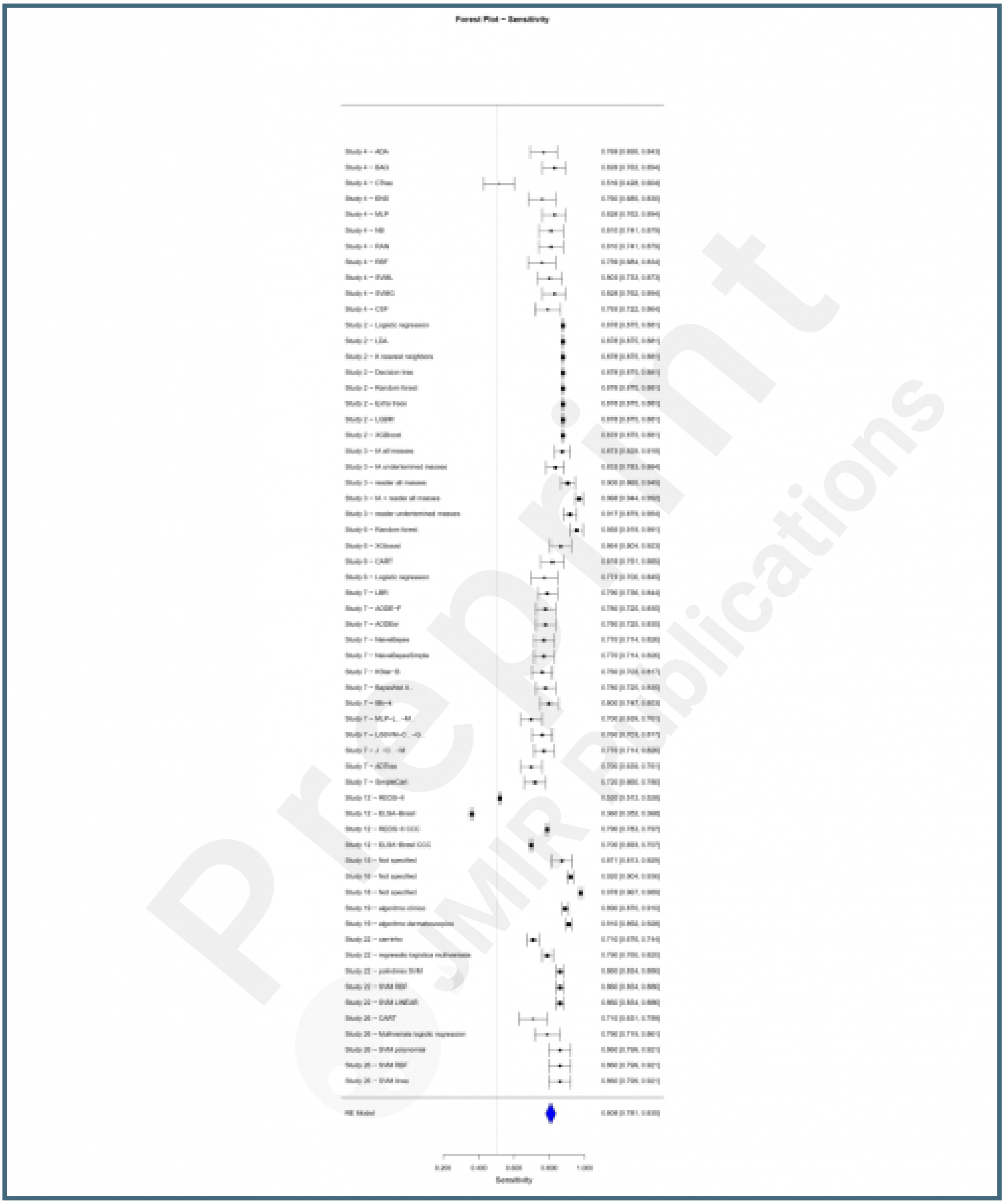
Figures

Study Selection.

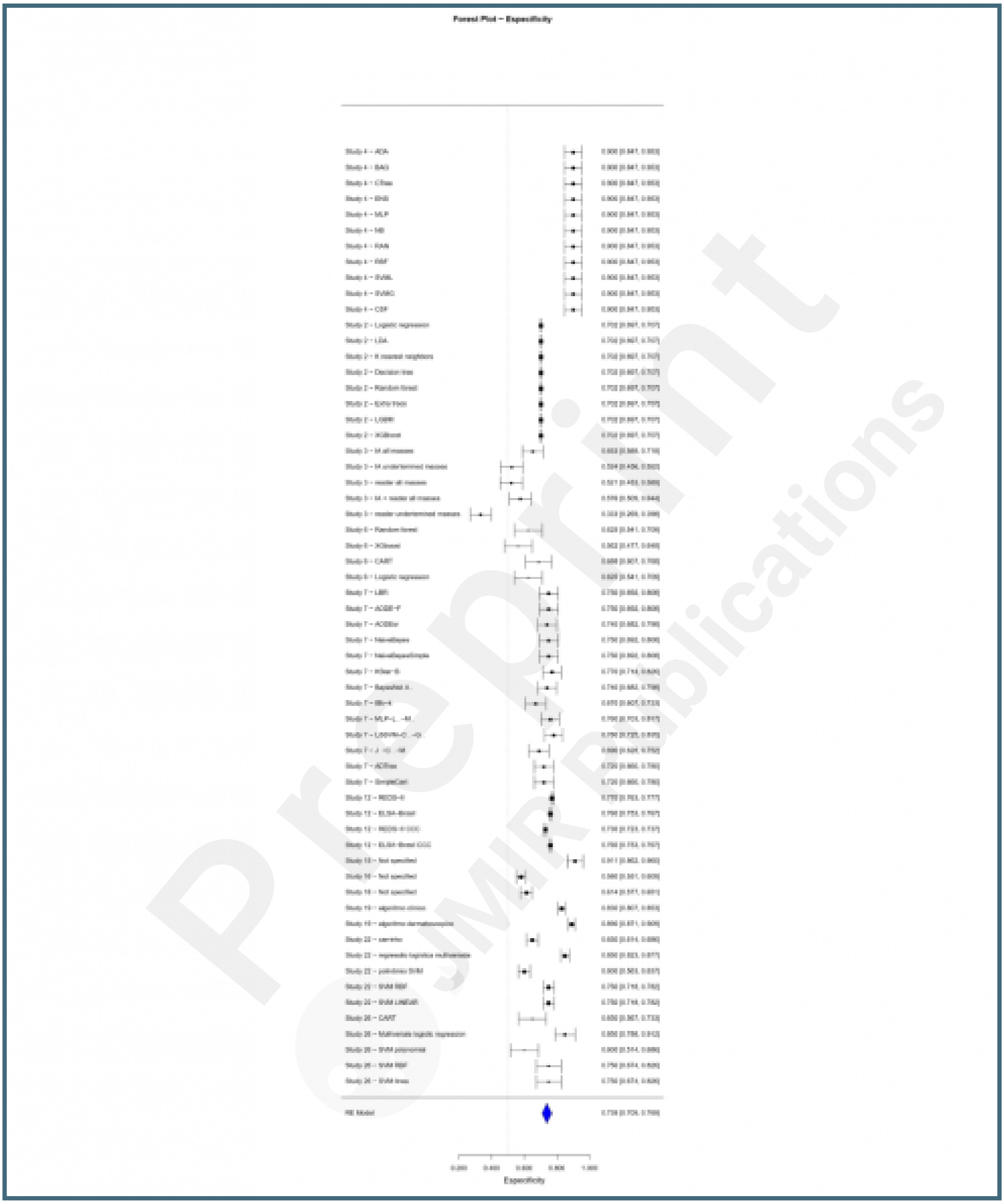


Risk of bias.

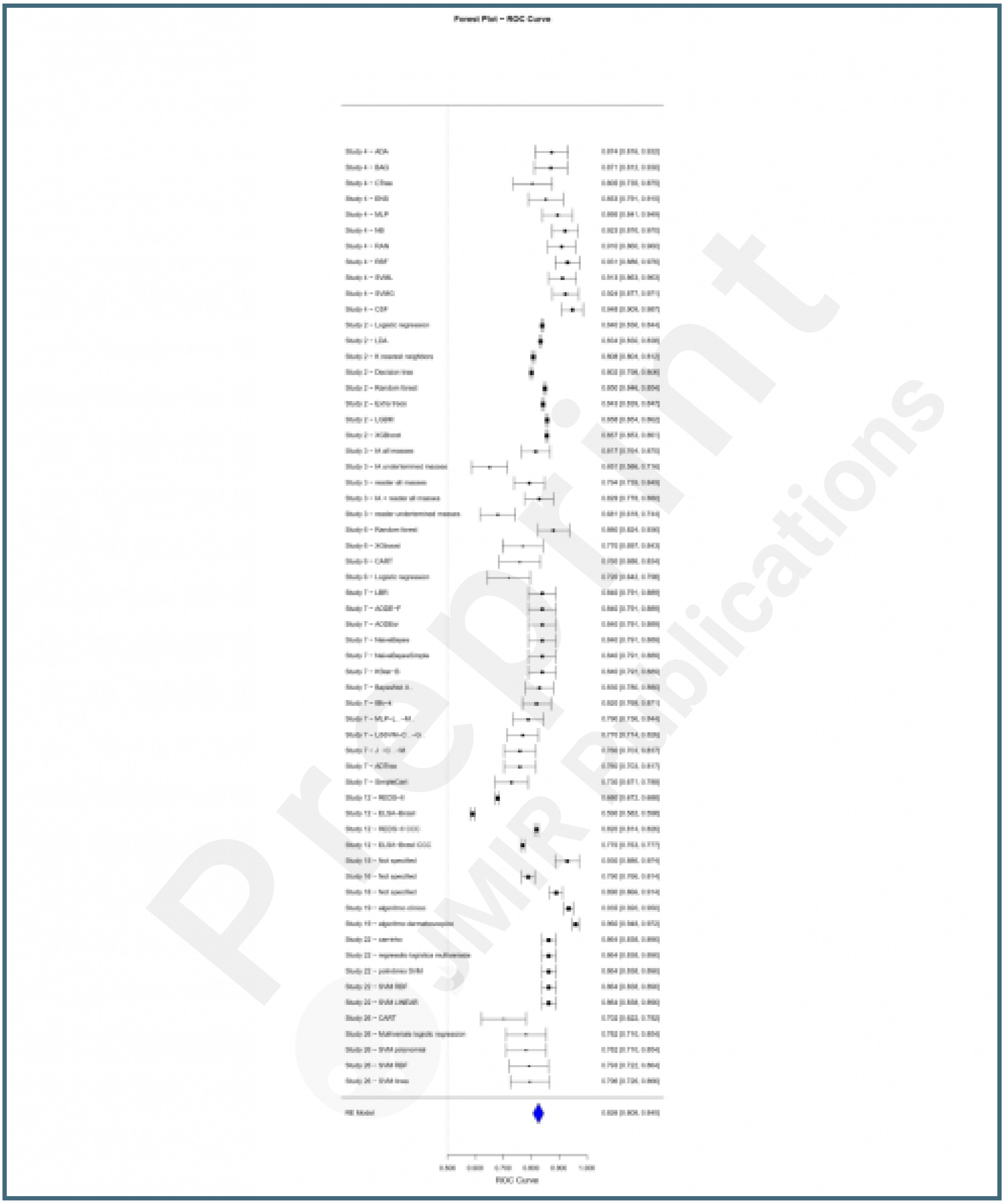
Forest Plot Sensitivity.



Forest Plot Especificity.



Forest Plot ROC Curve.



Preprint JMIR Publications

Multimedia Appendixes

Preprint JMIR Publications

String.
URL: http://asset.jmir.pub/assets/fb56a2ee7096d0466fcc86ddc43cd55a.pdf

Preprint JMIR Publications