Journal of Neonatal Surgery
ISSN (Online): 2226 -0439
Vol. 14, Issue 32s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Accuracy Of Spirometric Interpretation by Artificial Intelligence-Based Software in Comparison with Spirometric Interpretation by A Qualified Respiratory Physician in A Tertiary Care Center in Chengalpattu District

Rohita S[1]*, Aruna Shanmuganathan[2], A. Chitrakumar[3], Sruthi P. Mohan[4], J.Sam Selva Shruthi[5]

[1]Department of Respiratory Medicine, Karpaga Vinayaga Institute of Medical Sciences and Research Center. P.O, GST Road, Chinna Kolambakkam, Palayanoor, Maduranthakam- 603 308, Tamilnadu, India.

Orcid Id: 0009-0001-9883-2651

[2]Department of Respiratory Medicine, Karpaga Vinayaga Institute of Medical Sciences and Research Center. P.O, GST Road, Chinna Kolambakkam, Palayanoor, Maduranthakam- 603 308, Tamilnadu, India.

Orcid Id: 0000-0002-4942-6366

[3]Department of Respiratory Medicine, Karpaga Vinayaga Institute of Medical Sciences and Research Center. P.O, GST Road, Chinna Kolambakkam, Palayanoor, Maduranthakam- 603 308, Tamilnadu, India.

Orcid Id: 0000-0001-6775-0402

[4]Department of Respiratory Medicine, Karpaga Vinayaga Institute of Medical Sciences and Research Center. P.O, GST Road, Chinna Kolambakkam, Palayanoor, Maduranthakam- 603 308, Tamilnadu, India.

Orcid Id: 0009-0003-6271-4383

[5]Department of Respiratory Medicine, Karpaga Vinayaga Institute of Medical Sciences and Research Center. P.O, GST Road, Chinna Kolambakkam, Palayanoor, Maduranthakam- 603 308, Tamilnadu, India.

Orcid Id: 0009-0001-0608-6277

***Corresponding Author:**

Dr. Rohita S

*MD Postgraduate Department of Respiratory Medicine Karpaga Vinayaga Institute of Medical Sciences and Research Center P.O, GST Road, Chinna Kolambakkam, Palayanoor, Maduranthakam- 603 308, Tamil Nadu, India

Email ID: mailmer2.4.11@gmail.com

Orcid ID: 0009-0001-9883-2651

Cite this paper as: Rohita S, Aruna Shanmuganathan, A. Chitrakumar, Sruthi P. Mohan, J.Sam Selva Shruthi, (2025) Accuracy of Spirometric Interpretation by Artificial Intelligence-Based Software in Comparison with Spirometric Interpretation by A Qualified Respiratory Physician in A Tertiary Care Center in Chengalpattu District. *Journal of Neonatal Sccurgery*, 14 (32s), 810-817.

ABSTRACT

Background: Artificial Intelligence (AI) is transforming respiratory care by enhancing diagnostic accuracy and streamlining workflows. The efficacy of free AI tools for spirometry interpretation, particularly in the Indian population, remains largely unassessed. This study aimed to evaluate the diagnostic accuracy of ChatGPT for spirometry interpretation compared with that of qualified respiratory physicians in a south Indian tertiary care setting.

Methods: This cross-sectional study included 100 anonymised spirometry reports that met the ATS/ERS criteria. These reports were interpreted by respiratory physicians (gold standard), and the same reports were uploaded to ChatGPT. Interpretations of the spirometry were based on ATS/ERS guidelines using Z-scores and flow-volume loops. Statistical analyses included specificity, sensitivity using proportion agreement, kappa statistics, and ROC curve analyses.

Results: The 100 spirometry reports simultaneously analysed by AI & respiratory physicians overall normal vs. abnormal classification accuracy was 99%. For Z-score interpretation, the AI reported normal (25% vs. 26%), restriction (22% vs. 19%) obstruction (39% vs. 28%), and mixed (17% vs. 24%) compared to respiratory physician interpretation. In classifying flow-volume loop patterns, AI showed normal (30% vs 26%), restriction (18% vs. 26%) obstruction (50% vs. 38%), and mixed (2% vs. 10%) compared with respiratory physicians. In the final interpretation combining z-score & flow volume loop, AI interpretation was - 25% normal, 19% restriction, 39% obstruction, & 17% mixed, compared to the respiratory physician interpretation 26% normal, 21% restriction, 29% obstruction, & 24% mixed. AI achieved 99% agreement for normal, 98% for restriction, 90% for obstruction 93% for mixed.

Conclusion: ChatGPT is a promising tool for spirometry interpretation, but other similar AI platforms with larger samples need to be assessed before formal recommendations can be made.

Keywords: *Artificial Intelligence, Spirometry Interpretation, Diagnostic Accuracy, Pulmonologist, ChatGPT.*

1. INTRODUCTION

Artificial Intelligence (AI) is transforming respiratory care by enhancing diagnostic accuracy, enabling remote patient monitoring through wearable devices, and supporting personalised treatment plans. AI algorithms can analyse large datasets, including imaging and clinical parameters, to facilitate early disease detection and optimise therapy. The integration of AI, Machine Learning (ML), and Deep Learning (DL) into healthcare has demonstrated potential in diagnostic accuracy and clinical decision-making.[1] In respiratory medicine, AI applications span cancer screening, spirometry, polysomnography, and radiological interpretations such as X-rays and CT scans. Additionally, AI aids in real-time ventilator management, thereby improving outcomes in critically ill patients.[2] India, which accounts for approximately 18% of the global population, is currently experiencing an increasing burden of chronic respiratory diseases. However, a systematic and comprehensive analysis of their distribution and temporal trends across all Indian states remains insufficiently documented.[3]

Office spirometry is the most widely utilised, simple, and non-invasive pulmonary function test.[4] Although established guidelines exist for interpreting spirometry results, accurate analysis requires considerable expertise to evaluate normal values, flow-volume loops, and curves, and to clinically correlate them to reach an appropriate diagnosis [5]

Spirometry alone cannot establish a definitive clinical diagnosis without correlating the results clinically, highlighting an important limitation of the test.[6] AI holds the potential to significantly reduce the time required for interpreting spirometry traces from several minutes to mere seconds, thereby supporting healthcare professionals in both performing and interpreting these tests. Such rapid analysis can streamline clinical workflows, especially in high-volume or resource-limited settings.[7]

Despite these advantages, the adoption of AI in clinical trials and practice faces several challenges, including ethical concerns, high implementation costs, lack of algorithm transparency, and limited generalisability across diverse populations.[8] An ML model has demonstrated promising accuracy and precision (exceeding 90%) in identifying obstructive ventilatory patterns on spirometry among smokers in primary care who lack a prior respiratory diagnosis.[9] This model utilised pre-bronchodilator theoretical FEV1 values and underscored AI's potential in the early detection of obstructive airway diseases, such as COPD.[9]

Enhancing the clinical applicability and reliability of such models requires further research incorporating broader clinical parameters such as symptoms, imaging findings, and comorbidities and investigating practical approaches for ethical, cost-effective integration into routine workflows.[9] ML and DL algorithms have emerged as effective diagnostic tools, accurately classifying respiratory diseases such as pneumonia, fibrosis, cancer, tuberculosis, emphysema, and asthma using radiographic and CT imaging data.[10]

Although several free-to-use AI-based software tools are currently available, their efficacy in accurately interpreting spirometry results remains inadequately assessed, particularly in the Indian population. By addressing this gap, the present study evaluated the diagnostic accuracy of a freely available AI-based model, ChatGPT, for spirometry interpretation, compared with interpretations made by qualified respiratory physicians in a South Indian tertiary care setting.

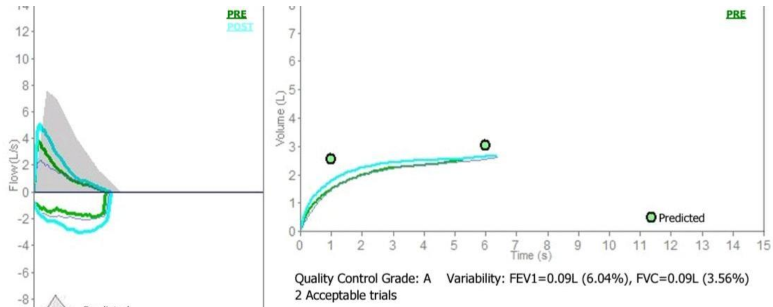
1. OBJECTIVES

To compare the accuracy of spirometry interpretation by AI-based software with that by qualified respiratory physicians in a tertiary care centre in the Chengalpattu district.

2. MATERIALS AND METHODS

This cross-sectional study was conducted in the Department of Respiratory Medicine at the Karpaga Vinayaga Institute of Medical Sciences and Research Centre. The study population comprised 100 spirometry reports that were interpreted by an expert qualified respiratory physicians and an AI-based software, ChatGPT. The study was initiated following ethical approval, with informed consent obtained, and patient confidentiality was maintained through the anonymisation of spirometry reports.

Figure 1: Spirometry Report



Interpretation

[redacted]

PRE Trial date 27-09-2023 12:21:28 **POST Bronchodilation with Salbutamol - 13:24:44**

Parameters		LLN	Pred	Best	%Pred	Z-score	PRE # 1	PRE # 2	PRE # 3	POST	%Pred	%Chg
FVC	L	1.99	3.05	2.62*	86	-0.67	2.62	2.53		2.69*	88	3
FEV1/FVC	%	74.3	84.5	60.3*	71	-3.91	56.9	**62.5**		69.5*	82	15
FEV1/FEV6	%		83.7	57.5	69		57.5	62.5		69.8	83	21
FEV1	L	1.68	2.55	1.58*	62	-1.85	1.49	**1.58**		1.87*	73	18
FEV6	L	1.99	3.05	2.59	85	-0.71	2.59	2.53		2.68	88	3
ELA	Years		41	83	202		83	79		68	166	-18
PEF	L/s	4.19	7.61	3.92*	51	-1.78	2.39	**3.92**		5.08*	67	30
FET	s		6.00	6.42	107		6.42	5.23		6.34	106	-1
FEF2575	L/s	1.48	3.26	0.78	24	-2.29	0.78	**0.94**		1.26	39	62

*Best values from all loops - BTPS 1.063 31 °C (87.8 °F) - Predicted Knudson

2.1 Inclusion and exclusion criteria

Spirometry reports of adult patients that fulfilled the acceptability and repeatability criteria according to the ATS/ERS guidelines were included, while those that did not meet these criteria were excluded.

2.2 Sample size calculation

A purposive sampling method was employed, and the sample size was calculated based on a previous study by Savic-Pesic et al. (2023), published in the Multidisciplinary Digital Publishing Institute, which reported an accuracy of 97.5% for the Espiro app when compared with pulmonologist interpretations as the gold standard. Assuming a 95% confidence interval, an absolute precision of 5%, and an available population size of 150, the minimum required sample size was estimated to be between 100 and 110, calculated using the following formula:

$$n = \frac{Z^2_{1-\frac{\alpha}{2}}\, p(1-p)}{d^2}$$

2.3 Methods

Spirometry reports fulfilling the ATS/ERS criteria were interpreted by respiratory physician to establish a "gold standard" reference. The same set of spirometry reports was uploaded to an AI-based software, ChatGPT, for independent interpretation. The accuracy of ChatGPT's interpretations was compared with the interpretations provided by respiratory physician.

Spirometry interpretations adhered to the American Thoracic Society (ATS) and European Respiratory Society (ERS) guidelines, utilising Z-scores and flow-volume loop configurations. The normal Z-score range is between -1.64 and +1.64, corresponding to the 5th and 95th percentiles of the population. The final interpretation reports classified spirometry findings as normal, obstructive, restrictive, or mixed based on specific criteria for FEV1/FVC and FVC, where a Z-score >-1.645 indicates a normal or preserved value and <-1.645 indicates an abnormal or reduced value for both the parameters.

Statistical Analysis

The primary variable of interest was spirometry interpretation (Normal/Obstructive/Restrictive/Mixed), which was categorical and summarised using frequency and percentage. Agreement between the AI and physician interpretations, considered either categorical or continuous, was assessed using proportion agreement and the Kappa statistic, with ROC curve analysis employed to further evaluate diagnostic performance.

Discrepancies between the AI and physician interpretations, also a categorical variable, were quantified by frequency and percentage of mismatches and subjected to Chi-square or Fisher's exact tests as appropriate.

3. RESULTS

The mean age of the patients was 43.26 ± 15.39 years, with a nearly equal gender distribution (53% male, 47% female).

Regarding pulmonary function, the mean forced vital capacity (FVC) z-score was –1.23 ± 1.22), while the mean FEV_1/FVC z-score was lower at –1.82 ± 1.80, suggesting a more pronounced deviation from normative values in the FEV_1/FVC ratio (Table 1).

Table 1: Demographic and spirometry profile of the study population

Age (years)	43.26 ± 15.39
Female	47(47%)
Male	53(53%)
FVC Z-score	-1.23 ± 1.22
FEV1/FVC Z-score	-1.82 ± 1.8

Of the 100 cases, the AI identified 75 as abnormal and 25 as normal. Among the abnormal predictions, 74 were true positives correctly identified by both the AI and the respiratory physician, with only 1 false positive, where the AI predicted abnormal, but the respiratory physician deemed the case normal. AI achieved 0 false negatives and 25 true negatives.

Overall, the AI achieved a sensitivity of 100%, reflecting its ability to detect all abnormal cases without missing any. The specificity was 96.15%, representing a minimal false-positive rate. The PPV was 98.67%, indicating that nearly all AI-predicted abnormal cases were truly abnormal, whereas the NPV was 100%, confirming that all AI-predicted normal cases were accurately identified (Table 2).

Table 2: Cross-tabulation of AI and Respiratory physician interpretations for binary classification of pulmonary function (normal vs. abnormal)

	Respiratory physician Abnormal	**Respiratory physician Normal**	**Total (AI Prediction)**
AI Abnormal	74	1	75
AI Normal	0	25	25
Total (Respiratory physician)	74	26	100

In the flow-volume loop pattern analysis, ChaGPT in comparison to respiratory physicians identified normal patterns (30% vs 26%), restrictive (26% vs. 18%), mixed (10% vs. 2%) and obstructive patterns (50% vs 38%)

Based on Z-Score interpretation, the percentage of normal interpretations remained nearly identical between the two groups (26% vs. 25%), whereas respiratory physician identified more restrictive (22% vs. 19%) and mixed (24% vs. 17%) cases. AI reported a higher prevalence of obstruction (39%) as compared to respiratory physician (28%),

In terms of the final interpretation based on both flow-volume loop and z-score the normal spirometry interpretation was consistent across both methods (26% vs. 25%). Respiratory physician assigned a higher proportion of cases to the mixed (24% vs. 17%) and restrictive (21% vs. 19%) categories. A greater number of cases classified as obstructive (39%) by ChatGPT as opposed to respiratory physician (29%). (Table 3).

Table 3: Comparison of pulmonologist and AI-based interpretations across loop pattern, z-score, and final diagnosis categories

Interpretation Type	**Pulmonologist N=100 (%)**	**AI N=100 (%)**

Loop Pattern	Normal	26 (26%)	30 (30%)
	Restriction	26 (26%)	18 (18%)
	Obstruction	38 (38%)	50 (50%)
	Mixed	10 (10%)	2 (2%)
Z-score	Normal	26 (26%)	25 (25%)
	Restriction	22 (22%)	19 (19%)
	Obstruction	28 (28%)	39 (39%)
	Mixed	24 (24%)	17 (17%)
Final Diagnosis	Normal	26 (26%)	25 (25%)
	Restriction	21 (21%)	19 (19%)
	Obstruction	29 (29%)	39 (39%)
	Mixed	24 (24%)	17 (17%)

In case of normal patterns, the AI achieved the highest diagnostic accuracy of 99%, with a sensitivity of 96.15%, 100% specificity, and PPV. The negative predictive value (NPV) was also high at 98.67%, with only a single misclassified case.

For restrictive patterns, the AI showed an accuracy of 98%, sensitivity of 90.78%, and 100% specificity and PPV. The NPV remained high at 97.53%.

Regarding obstructive patterns, the AI attained a 100% sensitivity and NPV of 100%, indicating no missed true-positive cases. However, the specificity (85.92%) and PPV (74.36%) were notably lower owing to 10 false positives, which reduced the overall accuracy to 90%

For mixed patterns, the AI demonstrated specificity and a positive predictive value (PPV) of 100%, indicating that all AI-identified mixed cases were true positives. However, the sensitivity was comparatively lower at 70.83%, resulting in the AI missing seven cases confirmed by pulmonologists. This resulted in an overall accuracy of 93%. (Tables 4 and 5).

Table 4: Concordance between AI-based and pulmonologist interpretations of ventilator patterns

AI Interpretation	**Pulmonologist Normal**		**Pulmonologist Restriction**		**Pulmonologist Obstruction**		**Pulmonologist Mixed**	
	Yes	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**
Normal	25	0	0	0	0	0	0	0
Abnormal	1	74	0	0	0	0	0	0
Restriction	0	0	19	0	0	0	0	0
Not Restriction	0	0	2	79	0	0	0	0
Obstruction	0	0	0	0	29	10	0	0
Not Obstruction	0	0	0	0	0	61	0	0
Mixed	0	0	0	0	0	0	17	0
Not Mixed	0	0	0	0	0	0	7	76

Table 5: Diagnostic performance metrics of AI-based interpretation for various patterns

Pattern	**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**	**Accuracy (%)**
Normal	96.15	100	100	98.67	99

Restriction	90.78	100	100	97.53	98
Obstruction	100	85.92	74.36	100	90
Mixed	70.83	100	100	91.57	93

4. DISCUSSION

In our study, the AI-based ChatGPT software of spirometry reports demonstrated 99% accuracy with the respiratory physician interpretation in identifying normal from abnormal. Using flow-volume loop and z-score-based interpretation, ChatGPT showed accuracy of 99%, 98% 93% and 90% for normal, restriction, mixed and obstruction, respectively. The AI interpretation of the obstructive pattern had least accuracy. This is because AI interpreted the lower FEV1 parameter (as air trapping) to be an obstructive pattern rather than mixed.

Our findings are supported by those reported by **Topalovic et al.,** who observed that pulmonologists' interpretations matched reference standards in 74.4% ± 5.9% of cases using ATS/ERS interpretative strategies. The AI-based software achieved 100% pattern matching and an 82% correct diagnostic categorisation rate, significantly outperforming pulmonologists with 44.6% accuracy ($p < 0.0001$).[11] In contrast, **Gottlieb et al.** showed AI with 95.6% sensitivity and 64.1% specificity for detecting pulmonary oedema, closely matching the physician's higher sensitivity of 96.7% but outperforming the specificity at 79.1%.[12] Similarly, **Saad et al.** showed that pulmonologists had difficulty diagnosing severity categories, with sensitivities of 47.73% for moderate, 44.01% for moderately severe, 60.34% for severe, and only 29.17% for others. The AI-based algorithm achieved an overall diagnostic accuracy of 86.59%.[13] **Das et al.** found pulmonologists showed greater agreement with correct AI interpretations ($p<0.001$), supporting AI as a decision-making aid.[14] **Sunjaya et al.** found that using a primary care spirometry dataset of 1113 patients, AI achieved 84.0% sensitivity, 86.8% specificity, and 85.4% overall accuracy for COPD detection, suggesting its utility in non-specialist settings.[15]

Our study has compared the AI vs expert interpretation as regards to the type of spirometric interpretation, as 26% normal, 22% restrictive, 24% mixed, and 28% obstructive whereas AI identified 25% normal, 19% restrictive, 17% mixed, and 39% obstructive. Among the abnormal AI had the highest accuracy in restriction followed by mixed and obstruction. A study by **Topalovic et al.** reported that z-score analysis effectively identified severe airflow obstruction in COPD and OAD, with FEV_1/FVC z-scores of –2.54 and –2.51, DLCO values of –2.77 and –1.89, and elevated RV (2.24) in OBD.[11] **Delclaux** highlighted that small airway obstruction with preserved FEV_1/FVC but reduced FEV_1 and VC presents diagnostic challenges. Our z-score-based AI interpretation reliably detected such patterns.[16] **Brusasco and Pellegrino** also emphasized that z-score-based interpretation enhances diagnostic precision in borderline cases and improves consistency.[17]

Furthermore, **Wang et al.** found that AI-assisted spirometry improved test quality over two months, with acceptability rates for FEV_1 and FVC increasing to 91.8% and 89.4%, respectively ($p < 0.0001$), and usability rates rising from 88% to over 99%.[18] **Moreno Mendez et al.** showed an AI model using pre-bronchodilator FEV_1 predicted values achieved 93% sensitivity, 97% specificity, and 94% precision, with 95% overall accuracy.[8]

Topole et al. showed that the ArtiQ.QC AI model achieved 87% accuracy, with 93% sensitivity but lower specificity (35%) across 8258 spirometry curves. Accuracy was highest in asthma (90%, specificity 84%) and lowest in COPD (43%) due to low specificity (52%) and PPV (25%).[19] **Robertson et al.** reported collaboration between pulmonologists and explainable AI improves interpretation accuracy, requiring appropriate training and oversight[20]

Our study confirms that ChatGPT spirometry interpretation aligns closely with respiratory physician assessments, particularly for normal, restrictive, and mixed patterns. Discrepancies in the classification of spirometry, AI demonstrated high diagnostic accuracy and consistency. These findings support the growing role of AI as a reliable adjunct in pulmonary diagnostics when integrated with clinical expertise.

5. LIMITATIONS

The generalisability of this study is constrained by its single-centre design and limited, homogeneous sample size, which may impact its external validity. Further concerns arise from the AI model's lack of specific training for spirometry interpretation and the omission of crucial clinical variables. When AI systems are trained on shared opinions of pulmonologists, they risk inheriting existing subjective biases which could be overcome by using a consensus expert spirometry interpretation. In view of several AI platforms available the need to validate their interpretations as well as to know the details of the training sets used are crucial in selecting the right tool.

6. CONCLUSION

The AI model, ChatGPT, showed 99 % accuracy and 100 % sensitivity for detecting normal versus abnormal spirometry. The obstructive pattern had the lowest (90%) accuracy in AI interpretation. This was due to overclassification into obstructive pattern by the AI model due to lower FEV1 attributed to air trapping. These results underscore the potential of AI tools as reliable adjuncts for pulmonary function interpretation, especially in high-volume, and limited availability of expert interpretation. ChatGPT is a promising tool for spirometry interpretation, but other similar AI platforms with larger training sets, incorporating clinical variables and multicentre data need to be assessed before formal recommendations can be made.

Disclaimer: Author's statement that the views expressed in the submitted article are his or her own and not an official position of the institution or funder.

Source of Funding: Nil

Conflict of interest: The authors declare that there is no conflict of interest.

Acknowledgements: The authors thank Dr.R.Annamalai, Dean and Managing Director, Karpaga Vinayaga Institute of Medical Sciences and Research Centre, India, for valuable support.

Consent to Publication: "Author(s) declared taking informed written consent for the publication of clinical photographs/material (if any used), from the legal guardian of the patient with an understanding that every effort will be made to conceal the identity of the patient, however it cannot be guaranteed."

REFERENCES

1. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J 2021;8: e188–94. https://doi.org/10.7861/fhj.2021-0095.
2. Al-Anazi S, Al-Omari A, Alanazi S, Marar A, Asad M, Alawaji F, et al. Artificial intelligence in respiratory care: Current scenario and future perspective. Ann Thorac Med 2024; 19:117–30. https://doi.org/10.4103/atm.atm_192_23.
3. India State-Level Disease Burden Initiative CRD Collaborators. The burden of chronic respiratory diseases and their heterogeneity across the states of India: The Global Burden of Disease Study 1990-2016. Lancet Glob Health 2018;6: e1363–74. https://doi.org/10.1016/S2214-109X(18)30409-1.
4. Ponce MC, Sankari A, Sharma S. Pulmonary function tests. Stat Pearls, Treasure Island (FL): Stat Pearls Publishing; 2025. https://www.ncbi.nlm.nih.gov/books/NBK482339/.
5. Sim YS, Lee J-H, Lee W-Y, Suh DI, Oh Y-M, Yoon J-S, et al. Spirometry and bronchodilator test. Tuberc Respir Dis (Seoul) 2017; 80:105–12. https://doi.org/10.4046/trd.2017.80.2.105.
6. Kaplan A, Cao H, FitzGerald JM, Iannotti N, Yang E, Kocks JWH, et al. Artificial intelligence/machine learning in respiratory medicine and potential role in asthma and COPD diagnosis. J Allergy Clin Immunol Pract 2021; 9:2255–61. https://doi.org/10.1016/j.jaip.2021.02.014.
7. Health Innovation Network, North-East and North Cumbria ArtiQ, a Clario company, Leuven, Belgium School of Psychology, Faculty of Health Sciences and Wellbeing, University of Sunderland. Using artificial intelligence to improve spirometry provision: a case study. Org.uk.2024 https://healthinnovationnenc.org.uk/wpcontent/uploads/2025/03/spirometry_final.pdf.
8. Krishnan G, Singh S, Pathania M, Gosavi S, Abhishek S, Parchani A, et al. Artificial intelligence in clinical medicine: catalyzing a sustainable global healthcare paradigm. Front Artif Intell 2023; 6:1227091. https://doi.org/10.3389/frai.2023.1227091.
9. Moreno Mendez R, Marín A, Ferrando JR, Rissi Castro G, Cepeda Madrigal S, Agostini G, et al. Artificial intelligence applied to forced spirometry in primary care. Open Respir Arch 2024; 6:100313. https://doi.org/10.1016/j.opresp.2024.100313.
10. Yadav P, Rastogi V, Yadav A, Parashar P. Artificial Intelligence: A promising tool in the diagnosis of respiratory diseases. Intelligent Pharmacy 2024; 2:784–91. https://doi.org/10.1016/j.ipha.2024.05.002
11. Topalovic M, Das N, Burgel P-R, Daenen M, Derom E, Haenebalcke C, et al. Artificial intelligence outperforms pulmonologists in the interpretation of pulmonary function tests. Eur Respir J 2019; 53:1801660. https://doi.org/10.1183/13993003.01660-2018.
12. Gottlieb M, Patel D, Viars M, Tsintolas J, Peksa GD, Bailitz J. Comparison of artificial intelligence versus real-time physician assessment of pulmonary oedema with lung ultrasound. Am J Emerg Med 2023; 70:109–12. https://doi.org/10.1016/j.ajem.2023.05.029.
13. Saad T, Pandey R, Padarya S, Singh P, Singh S, Mishra N. Application of artificial intelligence in the interpretation of pulmonary function tests. Cureus 2025;17: e82056. https://doi.org/10.7759/cureus.82056

14. Das N, Happaerts S, Gyselinck I, Staes M, Derom E, Brusselle G, et al. Collaboration between explainable artificial intelligence and pulmonologists improves the accuracy of pulmonary function test interpretation. Eur Respir J 2023;61. https://doi.org/10.1183/13993003.01720-2022.
15. Sunjaya A, Edwards GD, Harvey J, Sylvester K, Purvis J, Rutter M, et al. Validation of artificial intelligence spirometry diagnostic support software in primary care: a blinded diagnostic accuracy study. ERJ Open Res 2025:00116–2025. https://doi.org/10.1183/23120541.00116-2025.
16. Delclaux C. No need for pulmonologists to interpret pulmonary function tests. Eur Respir J 2019; 54:1900829. https://doi.org/10.1183/13993003.00829-2019.
17. Brusasco V, Pellegrino R. Pulmonary function interpretative strategies: from statistics to clinical practice. Eur Respir J 2022; 60:2200317. https://doi.org/10.1183/13993003.00317-2022.
18. Wang Y, Li Y, Chen W, Zhang C, Liang L, Huang R, et al. Deep learning for spirometry quality assurance with spirometric indices and curves. Respir Res 2022; 23:98. https://doi.org/10.1186/s12931-022-02014-9.
19. Topole E, Biondaro S, Montagna I, Corre S, Corradi M, Stanojevic S, et al. Artificial intelligence-based software facilitates spirometry quality control in asthma and COPD clinical trials. ERJ Open Res 2023; 9:00292–2022. https://doi.org/10.1183/23120541.00292-2022.
20. Robertson NM, Centner CS, Siddharthan T. Integrating artificial intelligence in the diagnosis of COPD globally: A way forward. Chronic Obstr Pulm Dis 2024; 11:114–20. https://doi.org/10.15326/jcopdf.2023.0449