

Using artificial intelligence to improve spirometry provision: a case study

Spirometry is key to diagnosing respiratory conditions such as asthma and COPD, but delivery in general practice has taken a severe hit since the COVID-19 pandemic. So could the use of artificial intelligence help to improve provision? This UK case-based study investigated the potential impact of AI-supported spirometry

CLAIRE ADAMS¹
ELENA SMETS²
JON REES³
JULIE MAES²
MARKO TOPALOVIC²

¹ Health Innovation Network, North-East and North Cumbria

² ArtiQ, a Clario company, Leuven, Belgium

³ School of Psychology, Faculty of Health Sciences and Wellbeing, University of Sunderland
 Contact: info@artiq.eu

Spirometry is a crucial tool in the diagnosis and management of respiratory conditions, such as asthma and chronic obstructive pulmonary disease (COPD).^{1,2} Specialist respiratory groups such as the American Thoracic Society and European Respiratory Society (ATS/ERS) have published guidelines on the execution and interpretation of spirometry.^{3,4} Specific standards have been published for primary care settings.⁵

The most convenient and cost-effective way of providing spirometry for patients is in primary care, but provision is suboptimal,

particularly following the COVID-19 pandemic.⁶ Studies show only 13.4% of spirometry performed in primary care fully meet ATS/ERS guidelines,⁷ and there are poor levels of agreement in spirometry interpretation between primary care staff and pulmonologists.⁶ Primary care practitioners report low confidence in identifying technical errors or interpreting spirometry.^{8,9} Important consequences include under-diagnosis, misdiagnosis and unnecessary referral to secondary care.

In the UK, spirometry can be provided by primary care clinicians in their surgeries, by referral to a locality-based service, or to a hospital lung function laboratory.¹⁰ With an aging population and an increasing prevalence of respiratory diseases, providing adequate spirometry services is essential to maintain the health and well-being of UK residents. However, workforce and equipment issues, lack of funding and challenges with guidelines and certification led to a crisis in spirometry provision, impacting patient care and outcomes.¹¹ Many services are redesigning their post-pandemic spirometry pathway to tackle these issues, with primary care networks (PCNs) introducing diagnostic hubs.⁹ The NHS Long Term Plan¹² has prioritised improving the quality and provision of spirometry, by providing training to primary care staff, i.e. through accreditation by the Association for Respiratory Technology and Physiology (ARTP). Asthma + Lung UK

THE STUDY

Methods: Two primary care clinicians participated in a 5-month trial. Spirometry sessions for 51 patients were recorded, and quality and interpretation with and without AI support were documented, together with the time taken for interpretation and confidence in the AI process.

Results: The quality of more than 90% of the spirometry sessions was evaluated by both the clinician and the AI to be good. The AI software matched the clinical diagnosis in 86% of the cases. In two cases where there was a mismatch, the clinician changed the suspected diagnosis to Interstitial Lung Disease (ILD) based on the AI feedback and referred to a chest clinic, where the diagnosis was confirmed. The average time to interpret the spirometry results decreased from 10.6 ± 4.1 min. to 5.6 ± 5.6 min ($p < 0.001$). The clinicians indicated on average they were very confident in the AI process.

Conclusion: This study shows that AI has the potential to reduce the time for interpretation of spirometry traces and support healthcare professionals in the execution and interpretation of spirometry. Routine AI use could enable conducting spirometry in primary care and reduce barriers such as workforce issues and lack of training. It could widen access to underserved populations and support accurate and early diagnosis of asthma, COPD, and other respiratory conditions such as ILD which are commonly mis- or under-diagnosed in practice.

suggests research is needed to identify the best ways to train staff and support them in interpreting results.¹¹

Previous studies have evaluated an artificial intelligence (AI) pulmonary function interpretation software and showed that it outperformed trained and trainee pulmonologists in the interpretation of hospital-based pulmonary function tests (PFTs).¹³ Recent studies have validated the AI solution (i.e. ArtiQ.Spiro) based on UK data,¹⁴ and showed an overall accuracy of 76% for detecting COPD, asthma and interstitial lung disease (ILD) based on spirometry data and basic demographic data. Additionally, the AI solution provides feedback on the quality of the spirometry traces according to ATS/ERS guidelines. Concordance between the AI algorithm and human over-reading was observed in 95% of the cases for FEV1, and in 91% of the cases for FVC.¹⁵

It can be hypothesised that AI-supported spirometry will improve spirometry pathways by improving quality of spirometry and the accuracy of diagnostic reports in primary care. It could reduce the time needed to interpret the spirometry results and provide more confidence to the clinician on the quality and interpretation of the spirometry session. As such, the introduction of AI-supported spirometry would lead to improved operational diagnostic efficiency (e.g. reduced dependence on secondary/tertiary care support) and improved service-user/provider experience (e.g. improved

confidence in spirometry usage). The current study aims to evaluate the impact of the ArtiQ.Spiro solution in a primary care setting.

METHODS

Participants: An experienced ARTP-accredited nurse and a primary care physician in the course of obtaining ARTP accreditation took part in the study in Sunderland. Both participants worked in their own surgeries.

Materials: Participants used SpiroConnect portable PC-based spirometers (MedChip, Chislehurst, Kent) with SpiroConnect software.

The SpiroConnect software is integrated with the AI solution (ArtiQ.Spiro, Leuven, Belgium), providing a quality and interpretation report. The quality report provides feedback on session level and trial-by-trial details based on ATS/ERS execution guidelines.³ The interpretation report provides feedback on the lung function interpretation according to ERS/ATS guidelines,⁴ as well as an AI-based disease probability, including asthma, COPD, ILD, normal and unidentified. The suggestions for follow-up are based on the suggested disease. Both reports can be accessed via the SpiroConnect software.

Study set-up: Table 1 describes the flow of the study. Participants were asked during a 5-month trial to include patients eligible for spirometry and document the reason for

referral (i.e. reason for executing spirometry) and which condition they clinically suspected. They first executed and interpreted the spirometry without AI support and documented the quality and interpretation of the session, together with the time taken to interpret the test and their confidence levels in the interpretation on a 5-point Likert scale. Next, they were asked to look at the AI software quality report and document if the AI confirms the acceptability of the session and whether this matched their own quality interpretation. Then they looked at the AI software interpretation report and documented if the physiological interpretation matched their own interpretation and whether they changed the diagnosis and/or management plan based on the provided disease probability and follow-up suggestions. They also documented the time taken to interpret the test using the AI software and rated their confidence in the AI process on a 5-point Likert scale.

Outcome measures: Outcomes include the quality analysis and interpretation of the spirometry session and the time taken to interpret the results with and without AI support. Additionally, the confidence of the clinicians in their interpretation and in the AI process was evaluated.

Statistical analysis: Analysis was performed under a frequentist framework using JASP 0.18.1 and with averages presented as means or medians as appropriate. Paired differences were investigated using a t-test for scale variables and a Wilcoxon test for non-parametric data. An alpha level of 0.05 for significance was used throughout.

RESULTS

51 patients were evaluated over the 5-month trial, equally provided by the two clinicians (25 and 26 patients, respectively). Table 2 provides an overview of the patient demographics and reasons for referral.

Quality evaluation of spirometry
47 sessions (92%) were evaluated by the clinician as valid. Only 1 session was found to be invalid (3 missing). AI rated 48 sessions as valid (94%) and 1 as invalid (2 missing). The AI quality assessment matched with the clinicians' assessment for 46 sessions (94%).

TABLE 1. DOCUMENTATION AND EVALUATION			
Clinical evaluation	Execute and interpret spirometry without AI support	Evaluate quality of spirometry with AI support	Evaluate interpretation of spirometry with AI support
If spirometry is needed, document: - Reason for referral - Suspected condition	Document: - If the session meets quality criteria - Interpretation of the session - Clinical diagnosis - Time taken to interpret the test - Confidence level in interpretation of the test (5-point Likert scale)	Document: - If AI confirms acceptability of session - If this matches own quality analysis	Document: - If the physiological interpretation matches own interpretation - If diagnosis changed based on AI report - If management plan changed based on AI report - Time taken to interpret the test - Confidence level in interpretation of the test (5-point Likert scale)

The results were not independently assessed by experts, so no ground truth was established, preventing the comparison of the AI and clinicians' assessments.

Interpretation of spirometry and diagnosis

Figure 1 presents on the left the distribution of physiological interpretations according to the clinicians, based on ATS/ERS guidelines, i.e. using z-scores (LLN) instead of a fixed ratio (70%) to define obstruction. On the right the distribution of clinical diagnoses based on patient history and spirometry interpretation (without support of AI) is shown. Main diagnoses include Normal (23), COPD (9) and asthma (7). Note that in several cases the physiological pattern was reported instead of an actual clinical diagnosis.

The AI suggested diagnosis matched the clinicians' diagnosis in 44 cases (86%). In 5 cases (10%) there was a mismatch (2 missing). In 2 of these 5 cases, the clinicians assessed the AI to be incorrect, based on clinical history of the patient, in 1 case the clinician's assessment was inconclusive and in 2 cases the clinician changed the diagnosis to likely ILD based on the AI suggested diagnosis of ILD. The AI feedback resulted in a change in follow-up investigations for 3 patients (6%) (e.g. check FeNo or refer to chest clinic for CXR and CT)

and resulted in a change of management for 5 patients (10%) (e.g. consider inhaled corticosteroid, change to LABA or refer to chest clinic).

Time and confidence assessment

The mean time to evaluate and interpret the spirometry session decreased significantly from 10.6±4.1 min. without AI to 5.6±5.6 min. with AI support ($p<0.001$). On average the clinicians rated their confidence in their diagnosis without AI support at 4±1 on a 5-point Likert scale (1 = not confident at all, 5 = extremely confident). They rated their confidence in the AI process at 3.9±1.3. Both time and confidence differed significantly between the two clinicians, with the more experienced, ARTP-accredited nurse being generally more confident and taking less time than the less experienced physician.

DISCUSSION

To our knowledge, this is the first study evaluating the impact of AI-supported spirometry in UK primary care. Results show both the clinicians and the AI rated more than 90% of the sessions as high quality. This is significantly better than average spirometry quality,⁷ as could be expected since clinicians were either ARTP-accredited or in the process of obtaining accreditation and therefore assumed to be better-trained than average primary care staff.

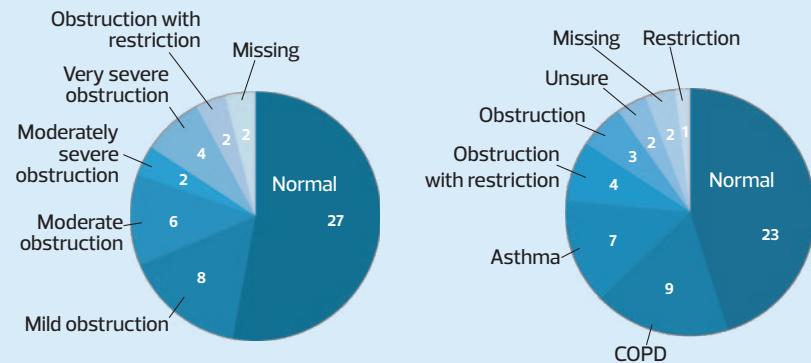
The clinicians' diagnosis was confirmed for the majority of the patients (86%). Although no reference diagnosis was obtained, this suggests a high diagnostic accuracy of the AI in line with earlier findings.¹⁴ Clinicians were on average very confident in their diagnosis, with the more experienced nurse being more confident than the less experienced physician, confirming that confidence and skills come with experience and training. In case of a mismatch, the clinicians indicated the AI to be too sensitive for asthma. In two cases the AI flagged ILD as a possible disease, which led to a referral to a chest clinic where the ILD diagnosis was confirmed. ILDs are a group of diffuse parenchymal lung disorders associated with substantial morbidity and mortality.¹⁶ Mortality rates can be as high as 80% over 5 years,¹⁷ so timely diagnosis is important for delaying disease progression and prolonging survival. However, opportunities for early diagnosis are often missed and research estimates diagnostic delays from 7 months to more than 5 years following onset of the symptoms.¹⁷ This case-study highlights the power of AI software to support clinicians in diagnosing severe respiratory diseases in an early stage. Other research using the same AI model supports these findings suggesting that the AI is capable of detecting ILD in 27% of the subjects 3.8±2.1 years prior to ILD diagnosis through standard care.¹⁸

The time to interpret the spirometry results was almost halved using the AI software and both clinicians were on average very confident in the AI process. This provides opportunities for future spirometry pathway improvements. To increase workforce capacity Asthma + Lung UK proposes to use a lower-band (3/4) member of staff, possibly through the Additional Roles Reimbursement Scheme (ARRS), who is trained to deliver the spirometry tests and a higher-band (6/7) colleague for interpretation or reporting.¹¹ This new pathway could widen the access to underserved populations. The AI software could complement this pathway to support the lower-band nurses with the execution of spirometry and the higher-band nurses with the interpretation, thereby reducing costs for execution and time for interpretation. The cost of a band 7 nurse for executing (45 min.) and interpreting spirometry (10 min.) is £29.49. The cost of an AI-supported band

TABLE 2. DEMOGRAPHICS AND REASON FOR REFERRAL

Age (mean±SD)	56±15 years	n=51
Gender	Male Female	19 30
Smoking status	Smokers Ex-smokers Non-smokers	14 16 18
Reason for spirometry	Known or suspected asthma Suspected COPD Cough (persistent) Other (e.g. previous abnormal spirometry, exertional wheeze, emphysema on CT)	9 8 19 13
Clinically suspected diagnosis	COPD Asthma Normal Unknown Long Covid	21 15 5 4 2

FIGURE 1. INTERPRETATION OF SPIROMETRY AND CLINICAL DIAGNOSES



Left, physiological interpretation of spirometry according to ERS/ATS guidelines (Z-scores); Right, clinical diagnosis based on spirometry and patient history without the support of AI

3 nurse to execute the spirometry (45 min. + £3 for the software) is £14.48, the physiological interpretation is done automatically by the AI. This leads to a cost reduction of £15.01 per spirometry. For the 51 patients in this study, this would equate to £765.51 of savings. Extrapolated to the North-East and North Cumbria region which performs around 84,000 tests a year (NENC internal data source), this would reduce costs from £2,477,160 to around £1,216,320,

resulting in a saving of £1,260,840 per year and could free up valuable nurse resources to deliver chronic disease management. Future research should investigate the feasibility of this pathway.

Limitations

This was a small case-study carried out with a particular cohort from Sunderland, therefore, the results may not be representative of the entire UK. The two

participating clinicians were highly trained in the execution and interpretation of spirometry compared with average GPs. This means the quality of the spirometry traces and match in diagnostic performance with the AI may not be representative of GPs in general. However, the fact that these highly trained clinicians trust the AI process provides optimism towards future adoption of the solution across a wider population. Additionally, it can be hypothesised that the impact of the AI could be even greater with less experienced physicians.

Conclusion

The AI software matched clinicians' spirometry quality evaluation and clinical diagnosis in the majority of the cases. Clinicians reported high levels of confidence in the AI process and a reduction in time to interpretation by almost 50%. Routine AI use could enable conducting spirometry in primary care and reduce barriers such as workforce issues and lack of training. It could widen access to underserved populations and support accurate and early diagnosis of asthma, COPD, and other respiratory conditions such as ILD which are commonly mis- or under-diagnosed in practice. ♦

REFERENCES

- Louis R, Satia I, Ojanguren I, et al. European Respiratory Society Guidelines for the Diagnosis of Asthma in Adults. *Eur Respir J* 2022; DOI: 10.1183/13993003.01585-2021
- NICE NG115. Chronic obstructive pulmonary disease in over 16s: diagnosis and management; 2019. <https://www.nice.org.uk/guidance/ng115>
- Graham BL, Steenbruggen I, Miller MR, et al. Standardization of Spirometry 2019 Update. An Official American Thoracic Society and European Respiratory Society Technical Statement. *Am J Respir Crit Care Med*. 2019; DOI: 10.1164/rccm.201908-1590ST
- Stanojevic S, Kaminsky DA, Miller MR, et al. ERS/ATS technical standard on interpretive strategies for routine lung function tests. *Eur Respir J*. 2022; DOI: 10.1183/13993003.01499-2021.
- Levy ML, Quanjer PH, Booker R, et al. Diagnostic spirometry in primary care: proposed standards for general practice compliant with American Thoracic Society and European Respiratory Society recommendations. *Prim Care Respir J* 2009; 18(3): 130–147.
- Doe G, Taylor SJ, Topalovic M, et al. Spirometry services in England post-pandemic and the potential role of AI support software: a qualitative study of challenges and opportunities. *Br J Gen Pract* 2023; DOI: 10.3399/bjgp.2022.0608.
- van de Hei SJ, Flokstra-de Blok BMJ, Baretta HJ, et al. Quality of spirometry and related diagnosis in primary care with a focus on clinical use. *NPJ Prim Care Respir Med*. 2020; DOI: 10.1038/s41533-020-0177-z.
- Dennis S, Reddel HK, Middleton S, et al. Barriers and outcomes of an evidence-based approach to diagnosis and management of chronic obstructive pulmonary disease (COPD) in Australia: a qualitative study. *Fam Pract* 2017;34(4):485–90. doi: 10.1093/fampra/cmw103
- Jenkins C. Spirometry performance in primary care: the problem, and possible solutions. *Prim Care Respir J* 2009;18(3):128–129
- White P, Wong W, Fleming T, Gray B. Primary care spirometry: test quality and the feasibility and usefulness of specialist reporting. *Br J Gen Pract* 2007;57(542):701–5
- Asthma + Lung UK. Diagnosing the problem: Right test, right time. <https://www.asthmaandlung.org.uk/diagnosing-problem-right-test-right-time-report>
- NHS England. The NHS Long Term Plan. 2019. <https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf>
- Topalovic M, Das N, Burgel PR, et al. Artificial intelligence outperforms pulmonologists in the interpretation of pulmonary function tests. *Eur Respir J*. 2019; DOI: 10.1183/13993003.01660-2018
- Elmahy A, Maes J, Desbordes P, et al. AI models for respiratory diseases from spirometry alone: validation based on UK Biobank. *Eur. Respir. J.* 2023, 62 (suppl 67) DOI: 10.1183/13993003.congress-2023.PA3508
- Cuyvers B, Desbordes P, Topole E, et al. AI Over-reading Based on ATS/ERS 2019 Criteria Is a Reliable Option for Instant Spirometry Quality Control in Clinical Trials (abstract). *Am J Respir Crit Care Med* 2023;207:A4064.
- Antoniou KM, Margaritopoulos GA, Tomassetti S, et al. European Respiratory Review 2014; DOI: 10.1183/09059180.00009113
- Case AH, Beegle S, Hotchkiss DL, et al. Defining the pathway to timely diagnosis and treatment of interstitial lung disease: a US Delphi survey. *BMJ Open Respiratory Research* 2023; DOI: 10.1136/bmjresp-2022-001594
- Topalovic M, Coenegrachts T, Van Steenberghe S, et al. Early detection of interstitial lung disease using AI-powered spirometry. *Eur. Respir. J.* 2022, 60 (suppl 66); DOI: 10.1183/13993003.congress-2022.1217