Allergy

WILEY | EAACI

LETTER OPEN ACCESS

Real-World Evaluation of AI-Assisted Readout of Skin Prick Automated Test Results

Julie van Waterschoot[1,2] | Sven F. Seys[3,4] | Ilan Baron[5] | Gregory Clement[6] | Julie Craps[7] | Tessa Dieudonné[8] | Glynnis De Greve[9] | Joris De Medts[10] | Natascha Friese[11] | Jan Hagemann[12] | Farah Hannachi[13] | Sofie Happaerts[14] | Edwin Hill[15] | Valérie Hox[16] | Bruno Lantsoght[17] | Winde Lemmens[18] | Patrick Levie[19] | Christophe Libeer[20] | Yves Mentens[21] | Tim S. Nawrot[22] | Haïg Nigolian[23] | Richard Oei[24] | Bas op de Coul[25] | Maike Pincus[26] | Mahadi Salah[27] | Christian Schellemans[28] | Salem Tamim[29] | Olivier Vanderveken[30] | Klara Van Gool[31] | Valérie Verkest[32] | Anne-Sophie Vinck[33] | Audrey Weinquin[34] | Rembert Daems[3,35] | Dirk Loeckx[3] | Senne Gorris[3,36] | Adam M. Chaker[37] | Laura Van Gerven[1,2,38]

Correspondence: Sven F. Seys (sven.seys@hippo-dx.com)

Received: 23 March 2026 | **Revised:** 9 June 2026 | **Accepted:** 22 June 2026

To the Editor,

Allergy diagnostics has taken an innovative turn with the introduction of the skin prick automated test (SPAT) [1, 2]. For respiratory allergy, skin prick testing (SPT) and serum specific IgE (sIgE) measurement are the gold standard methods to detect IgE-mediated sensitisation [3]. Despite its high sensitivity, manual SPT is subject to operator-dependent variability and prone to human errors. SPAT addresses these issues by standardizing the entire procedure, applying a fixed amount of allergen and a controlled prick force, in combination with AI-assisted readout to improve efficiency [1].

Previous research has demonstrated that SPAT, compared to manual SPT, shows lower intra-subject variability [1], higher consistency, and reduced patient discomfort [4]. When applying the validated 4.5-mm cut-off, SPAT shows equivalent diagnostic accuracy in detecting birch pollen and house dust mite allergies [2].

More recently, AI has been integrated into SPAT to provide wheal measurement suggestions, further increasing standardization and efficiency. Seys et al. described how the AI algorithm was trained ($n = 651$), validated ($n = 217$), and independently tested ($n = 95$) [5]. A strong correlation was observed between the physician-measured and the AI-measured longest wheal diameter (Pearson $r = 0.83$; $p < 0.0001$), with 5.8% of AI measurements adjusted by physicians, resulting in a change in test interpretation in 0.5% of cases.

To evaluate external validity in clinical practice, real-world SPAT data from 37 hospitals across 5 European countries were assessed. This cohort comprises test results from 10,756 patients (126,526 pricks) with respiratory (96.7%), food (3.3%), insect venom (0.06%) and drug (0.01%) allergens. AI-measured longest wheal diameters were compared with physician-verified AI measurements, confirming a strong correlation (Pearson $r = 0.96$, $p < 0.0001$; Figure 1). No wheal was detected by AI in 7.9% (10,031) of pricks, with only 0.3% (380) of pricks corresponding to a physician measurement of ≥ 4.5 mm. In total, 6.1% of the wheals were enlarged by the physician (median (IQR): +1.0 mm (+0.5; +1.6)), and 2.9% reduced (median (IQR): −0.8 mm (−1.8; −0.4)), resulting in a change in test interpretation—from negative to positive or vice versa—in 1.7% and 0.4% of the cases, respectively. An analysis of potential confounders is provided in the Supporting Information (Table S3–S5 and Figure S3).

This large real-world analysis demonstrates that the AI model is broadly generalizable and shows acceptable clinical performance. The low rate of interpretation-affecting adjustments is reassuring. The observed high correlation between AI and physician-verified AI measurements suggests that the AI model performs consistently across different hospitals. This is particularly relevant in comparison to manual SPT, where inter- and intra-observer variability has historically been a major limitation [6]. We observed that the total number of physician adjusted wheals (9.0%) in the current study was lower than the inter-observer readout variability (median: 19.8%) reported previously

For affiliations refer to page 3002.

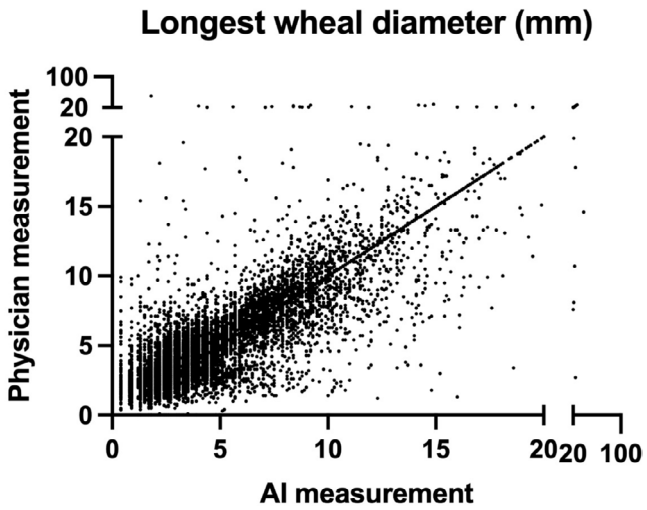

FIGURE 1 | Scatterplot of physician's and AI measurement of longest wheal diameter (116,495 pricks). Pearson $r = 0.96$ ($p < 0.0001$).

[5]. Hence, the physician and AI agree more often than different physicians do. While blinding physicians to AI output would ideally minimize anchoring bias, the real-world design of this study reflects routine clinical practice, where AI measurements are intentionally used to inform physician review rather than replace independent assessment.

These results support safe integration of the AI-assisted readout model into daily practice as a decision-support tool. Rather than replacing clinical judgment, the AI model functions as a reliable first reader, allowing physicians to focus on cases that require expert review [5]. By continuing use across an increasing number of patient populations, we aim to further improve performance, particularly regarding skin type variations where slightly lower performance was observed (Table S4 and Roux et al. [7]).

In conclusion, SPAT device and its AI-assisted readout method reduce variability, improve consistency, and increase time efficiency compared with manual SPT. This real-world evidence supports implementation in routine clinical practice, with clear benefits for physicians, healthcare systems, and patient care.

Author Contributions

S.F.S., S.G., and L.V.G. conceived the study, designed the data collection protocol, and performed the primary data analysis. J.W. and S.F.S. drafted the manuscript. S.F.S., D.L., S.G., J.W., and L.V.G. contributed to data acquisition and interpretation of results. D.L. and S.F.S. provided statistical support and reviewed the analytical approach. I.B., G.C., J.C., T.D., G.D.G., J.D.M., N.F., J.H., F.H., S.H., E.H., V.H., B.L., W.L., P.L., C.L., Y.M., T.S.N., H.N., R.O., B.C., M.P., M.S., C.S., S.T., O.V., K.V.G., V.V., A.-S.V., A.W., R.D., S.G., A.M.C., and L.V.G. critically revised the manuscript for important intellectual content. All authors contributed to and approved the final version of the manuscript.

Affiliations

[1]Allergy and Clinical Immunology Research Group, Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium | [2]Department of Otorhinolaryngology-Head and Neck Surgery, UZ Leuven, Leuven, Belgium | [3]Department of Otolaryngology, Head and Neck Surgery, Medical University of Vienna, Vienna, Austria | [4]Hippo Dx, Hasselt, Belgium | [5]Department of Otorhinolaryngology-Head and Neck Surgery, AZ Jan Portaels, Vilvoorde, Belgium | [6]Otorhinolaryngology-Head and Neck Surgery, Clinique Royale, Oostende, Belgium | [7]Department of Pneumology, AZ HH Lier, Lier, Belgium | [8]Department of Pneumology, AZ HH Mol, Mol, Belgium | [9]Department of Otorhinolaryngology-Head and Neck Surgery, ZAS Sint-Augustinus, Antwerp, Belgium | [10]Department of Otorhinolaryngology-Head and Neck Surgery, AZ Delta, Roeselare, Belgium | [11]HNO Praxis, Weil der Stadt, Germany | [12]Center for Rhinology and Allergology, Wiesbaden, Germany | [13]Immunology-Allergology Unit, Centre Hospitalier Luxembourg, Luxembourg, Luxembourg | [14]Department of Pneumology, AZ Klina, Brasschaat, Belgium | [15]Department of Pneumology, AZ Glorieux, Ronse, Belgium | [16]Service d'Otorhinolaryngologie, Cliniques Universitaires Saint-Luc, Brussels, Belgium | [17]NKO Dr Lantsoght, Knokke, Belgium | [18]Department of Otorhinolaryngology-Head and Neck Surgery, ZOL, Genk, Belgium | [19]Cabinet ORL Messidor, Brussels, Belgium | [20]Department of Pneumology, AZ Sint-Maarten, Mechelen, Belgium | [21]Department of Pneumology, AZ Herentals, Herentals, Belgium | [22]Center for Environmental Sciences, Hasselt University, Diepenbeek, Belgium | [23]Service D'allergology, Genève, Switzerland | [24]Department of Allergology, DC Klinieken, Groningen, the Netherlands | [25]Otorhinolaryngology, KNO Medisch Centrum, Waardenburg, the Netherlands | [26]Kinderartzpraxis Karow, Berlin, Germany | [27]Department of Otorhinolaryngology-Head and Neck Surgery, AZ Sint-Elisabeth, Zottegem, Belgium | [28]Doktershuis Antverpia, Antwerp, Belgium | [29]Service d'Allergology, Vivalia, Arlon, Belgium | [30]Department of Otorhinolaryngology-Head and Neck Surgery, UZA, Antwerp, Belgium | [31]NKO Kalmthout, Kalmthout, Belgium | [32]Department of Otorhinolaryngology-Head and Neck Surgery, ZAS Sint-Vincentius, Antwerp, Belgium | [33]Department of Otorhinolaryngology-Head and Neck Surgery, AZ Sint-Jan, Brugge, Belgium | [34]Cabinet ORL, Centre Hospitalier du Nord, Ettelbruck, Luxembourg | [35]IDLAB, Ghent University-imec, Ghent, Belgium | [36]Department of Otorhinolaryngology and Head and Neck Surgery, AZ Herentals, Herentals, Belgium | [37]Department of Otorhinolaryngology and Center of Allergy and Environment, TUM School of Medicine and Health, TUM University Hospital Klinikum rechts der Isar, Technical University of Munich, Munich, Germany | [38]Laboratory of Experimental Otorhinolaryngology, Department of Neurosciences, KU Leuven, Leuven, Belgium

Acknowledgements

Open Access funding provided by Medizinische Universitat Wien.

Funding

This work was supported by Hippo Dx.

Conflicts of Interest

S.F.S., R.D., D.L. and S.G. are employees of Hippocreates BV. S.F.S., R.D., D.L., S.G., L.V.G., and I.B. hold shares of Hippocreates BV. AMC reports grants, speaker honoraria, consultancy or advisory fees and/or research support and other, all via Technical University of Munich from Allergopharma, ALK Abello, Astra Zeneca, Bencard/Allergen Therapeutics, GSK, Novartis, Hippo Dx, LETI, Roche, Zeller, Sanofi, Regeneron, Thermo Fisher, European Institute of Technology (EIT Health) and Federal Ministry of Research and Education Germany. Other authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. S. Gorris, S. Uyttebroek, W. Backaert, et al., "Reduced Intra-Subject Variability of an Automated Skin Prick Test Device Compared to a Manual Test," *Allergy* 78, no. 5 (2023): 1366–1368.

2. S. F. Seys, A. Gherasim, F. Odul, et al., "Validation of the Skin Prick Automated Test (SPAT) Cut-Off Value in Birch Pollen and House Dust Mite Allergic Rhinitis Patients," *Allergy* 80, no. 12 (2025): 3302–3309.

3. T. Gureczny, B. Heindl, L. Klug, F. Wantke, W. Hemmer, and S. Wöhrl, "Allergy Screening With Extract-Based Skin Prick Tests Demonstrates Higher Sensitivity Over In Vitro Molecular Allergy Testing," *Clinical and Translational Allergy* 13, no. 2 (2023): e12220.

4. S. F. Seys, K. Roux, C. Claes, et al., "Skin Prick Automated Test Device Offers More Reliable Allergy Test Results Compared to a Manual Skin Prick Test," *Rhinology Journal* 62, no. 2 (2024): 216–222.

5. S. F. Seys, V. Hox, A. M. Chaker, et al., "Artificial Intelligence (AI)-Assisted Readout Method for the Evaluation of Skin Prick Automated Test Results," *Nature Communications* 16, no. 1 (2025): 8637.

6. W. A. McCann and D. R. Ownby, "The Reproducibility of the Allergy Skin Test Scoring and Interpretation by Board-Certified/Board-Eligible Allergists," *Annals of Allergy, Asthma & Immunology* 89, no. 4 (2002): 368–371.

7. K. Roux, S. F. Seys, V. Hox, et al., "Impact of Real-World Confounders on the Accuracy of an AI Model to Support Read Out of Skin Prick Automated Test Results," *Rhinology*, ahead of print, June 3, 2026, https://doi.org/10.4193/Rhin25.634.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Table S1:** Overview of tested allergens. **Table S2:** Overview of hospitals per country. **Table S3:** Impact of sex. **Table S4:** Impact of skin type. **Table S5:** Impact of allergen type. **Table S6:** Patient characteristics. **Figure S1:** Representative composite image with measurement of the longest wheal diameter. **Figure S2:** Bland–Altman plot of AI versus physician-verified AI measurement. **Figure S3:** Accuracy versus number of tests stratified per hospital.