Diagnosing spondyloarthritis by means of an exhaled breath analysis

By:
Manouk Vrielink
S2627477
Master Thesis Health Sciences
Science and Technology Faculty (TNW), University of Twente
Medisch Spectrum Twente, rheumatology department

First supervisor:
Dr. P.M. ten Klooster
University of Twente

Second supervisor:
Prof. Dr. J.A.M. van der Palen
University of Twente, Medisch Spectrum Twente

External supervisor:
Prof. Dr. H.E. Vonkeman
University of Twente, Medisch Spectrum Twente

Daily supervisor:
D. Gerritsen
Medisch Spectrum Twente

28-06-2024

Abstract

Introduction: Early diagnosis and treatment of spondyloarthritis (SpA) is crucial for preventing structural damage and stiffness often encountered in advanced stages of the disease. Currently, diagnosis relies on identifying typical SpA features along with laboratory findings and imaging techniques. Nonetheless, diagnostic challenges exist due to the absence of diagnostic criteria and the heterogeneous presentation of SpA. Exhaled breath analysis of volatile organic compounds (VOCs) through an electronic nose offers a promising opportunity for early, non-invasive diagnosis. This technique has already demonstrated effectiveness in identifying various conditions. This study aimed to investigate the ability of the electronic nose to differentiate between healthy controls and SpA patients by means of an exhaled breath analysis.

Methods: Data were collected between December 2021 and April 2024. All participants performed a breath test with the aeoNose and provided relevant demographic and medical information. Random forest machine learning with 10-fold cross-validation was employed to analyze the breath samples and create an optimal discriminating algorithm. The performance of this model was compared to that of a multivariate logistic regression model, which incorporated readily available clinical variables in addition to the aeoNose predictive values.

Results: Breath samples of 59 SpA patients (mean±SD age 50.3±15.8 years, 61% male) were compared with 180 healthy controls (mean±SD age 53.3±17.6 years, 42.8% male). The model that only included the aeoNose classification value resulted in an area under the receiver operator curve (AUC-ROC) of 0.95 (95% CI 0.92-0.99), a sensitivity of 95%, and a negative predictive value (NPV) of 98%. The second model, including solely the clinical variables age and gender, resulted in an AUC-ROC of 0.63 (95% CI 0.54-0.71), a sensitivity of 95%, and an NPV of 85%. Combining the clinical variables with the aeoNose classification value in a multivariate logistic regression model only slightly improved overall performance with an AUC-ROC of 0.96 (95% CI 0.93-0.99), while sensitivity remained 95% and NPV 98%.

Conclusion: The aeoNose shows promise in discriminating between SpA patients and controls with high diagnostic accuracy, indicating its potential use as a diagnostic tool. However, cross-validation of the algorithm in independent samples is necessary. Adding readily available clinical variables in a multivariate logistic regression model only slightly enhances performance.

Introduction

Spondyloarthritis (SpA) encompasses a heterogeneous group of musculoskeletal inflammatory diseases that share common clinical features, genetic susceptibility, and pathophysiological mechanisms [1]. SpA mainly affects young adults, with symptoms typically emerging between the ages of 20 and 40 [2]. Based on the Assessment in Spondyloarthritis International Society (ASAS) classification criteria and clinical presentation, SpA can be categorized into two primary subtypes [1]. Axial SpA (axSpA) predominantly affects the axial skeleton, including the spine and sacroiliac joints, whereas peripheral SpA (pSpA) is characterized by arthritis, enthesis, and/or dactylitis [1, 3]. Additionally, axSpA can be further categorized into two subgroups: axSpA with radiographic sacroiliitis, also referred to as ankylosing spondylitis (AS), and axSpA without radiographic sacroiliitis [4]. In this differentiation, the term radiographic sacroiliitis refers to erosions and/or ankylosis seen on classic X-rays of the sacroiliac joints. In early phases of spondyloarthritis, magnetic resonance imaging (MRI) of the sacroiliac joints may show signs of inflammation while the X-rays are still normal. Therefore, non-radiographic axSpA is often considered to be an early phase of radiographic axSpA [5-7]. Inflammatory lower back pain, morning stiffness, and sore joints are some of the most common complaints with SpA. Patients may also present with extra-musculoskeletal manifestations including uveitis, inflammatory bowel diseases, and psoriasis [8]. Although the exact pathogenesis of SpA is still unknown, multiple studies have shown the significant involvement of the human leukocyte antigen (HLA)-B27 in disease initiation, particularly in AS [9, 10]. Moreover, there is emerging evidence indicating the contribution of the intestinal microbiome to disease initiation [4]. In the Netherlands, the exact prevalence of SpA is currently unknown, with considerable variation in the reported prevalence rates across studies [1]. In Europe, the estimated prevalence of SpA ranges from 0.5% to 1.0% of the general population [11].

The diagnosis of SpA involves a clinical judgment based on both existing and absent features, while also considering alternative diagnoses. Typical SpA features include inflammatory back pain, dactylitis, asymmetrical arthritis, elevated acute-phase reactants, enthesitis, inflammatory bowel diseases, family history of SpA, marked response to NSAIDs, psoriasis, and uveitis [12, 13]. Additional laboratory investigations such as HLA-B27 status and levels of C-reactive protein (CRP) may aid in diagnosis, along with imaging of the sacroiliac joints [14]. Diagnosing SpA remains challenging, however, due to its heterogeneous presentation. No single feature or additional marker has sufficient specificity to decide on a diagnosis. Especially in early phases of the disease, diagnosis proves to be problematic in practice. Proposed imaging methods aimed at diagnosing early stages of axSpA also exhibit low specificity, as similar abnormalities can be observed in healthy runners, post-partum females, or those experiencing other traumas [15]. Therefore, diagnosis involves recognizing a pattern of features that collectively provide sufficient likelihood to diagnose the patient [16].

Studies have shown that on average, the time between symptom onset and the definitive diagnosis of SpA ranges from 5 to 12 years [2, 17, 18]. This very large delay in diagnosis has significant implications for disease outcomes and work productivity. Given that SpA affects mainly young individuals, acknowledging its impact on work productivity is crucial. Diagnostic delays longer than 5 years are associated with worsened disease outcomes, including loss of spinal range of motion and irreversible structural damage [13]. Moreover, indices measuring disease activity and functional impairment tend to worsen with longer diagnostic delays. Establishing the diagnosis early on allows the initiation of effective treatment, thereby preventing progressive structural damage and stiffness commonly observed in advanced stages of SpA [19, 20].

At present, SpA is increasingly treated using a Treat-to-Target (T2T) approach, which states that treatment strategies should target specific objectives, typically including sustained remission or, if not possible, low disease activity as measured by appropriate disease activity composite scores [21]. Treatment can involve pharmacological or non-pharmacological therapies, including physical exercise and physiotherapy. Nonsteroidal anti-inflammatory drugs (NSAIDs) are the first-line pharmacological therapy for decreasing inflammation, pain, and stiffness. Biological disease-modifying antirheumatic drugs (bDMARDs), and especially tumor necrosis factor inhibitors (TNFi), are reported as second-line treatment. Furthermore, in pSpA, methotrexate is used as a first-line conventional synthetic disease-modifying antirheumatic drug (csDMARD) [4, 13].

Considering the importance of early treatment, timely diagnosis of SpA is essential. Recently, multiple studies have shown the potential of volatile organic compounds (VOCs) in noninvasive, early diagnosis of various conditions, including different types of cancer and infectious diseases [22]. VOCs, present in exhaled breath, reflect both physiological and pathological processes within the human body, such as inflammation, oxidation, and infection. The premise is that metabolic and biochemical processes associated with various pathological states give rise to different endogenous VOCs [23]. These compounds form unique patterns, known as breath prints, and can be detected with chemical sensors. Given the inflammatory character of SpA, it is plausible that VOCs representing inflammation and activity of the disease can be found in exhaled breath. Previous studies demonstrated the effectiveness of utilizing VOCs in discriminating between patients with rheumatoid arthritis (RA) and healthy individuals. For instance, a diagnostic study by Brekelmans et al. (2016) illustrated that breath prints of RA patients could be distinguished from those of healthy controls with a sensitivity of 76% and a specificity of 67% [22]. These promising results require further investigation into the possibilities that VOCs offer in early diagnosis of other inflammatory rheumatic diseases.

The aeoNose electronic nose (the eNose company, Zutphen, the Netherlands) is a handheld device that can measure VOC patterns. To facilitate the recognition of various VOC mixtures, the aeoNose requires calibration [23]. This involves training the model with a large amount of breath prints, from which a database must be developed. Subsequently, newly detected scent patterns can be matched with these existing breath profiles using comparative pattern recognition analysis. This technique has already demonstrated effectiveness in diagnosing a variety of diseases. For instance, Kort et al. (2022) evaluated the aeoNose as a diagnostic tool for non-small cell lung cancer, yielding a sensitivity of 95% and a specificity of 49% in the external validation cohort [24]. Similarly, a study by van de Goor et al. (2018) investigated the potential of the aeoNose as a screening tool for lung cancer, achieving a sensitivity of 88% and a specificity of 86% in the blind test set [25]. Moreover, the diagnostic value of the aeoNose has been explored in numerous pilot studies involving other diseases, although these results have yet to be validated.

This pilot study describes the first step in investigating the ability of the aeoNose to distinguish between SpA patients and healthy individuals by means of an exhaled breath analysis. This concerns the learning phase, in which a robust prediction model must be developed and trained. A secondary aim is to explore whether the performance of the generated model can be optimized by adding relevant clinical variables.

Materials and method

Study design

The study was designed as a prospective, observational, and mono-centered diagnostic study at Medisch Spectrum Twente hospital, Enschede, the Netherlands. Data collection started in December 2021 and was continued until the end of April 2024. The study protocol was approved by the Institutional Review Board of Medisch Spectrum Twente (number K20-03).

Study population

SpA patients with scheduled appointments at the outpatient clinic of the rheumatology department were approached in the waiting room by the researcher. A brief overview of the procedure was provided, followed by an invitation to participate in the study. Those who agreed were then given the informed consent form (Appendix A). After their appointment with the rheumatologist, participants completed the breath test. The control group of people without a rheumatic disease was mainly collected during an earlier study, including mostly people who work at Medisch Spectrum Twente and family, friends, or caretakers of patients of the rheumatology department (D. Gerritsen, 2023, personal communication).

All subjects included in this study had to be 18 years or older and sign the informed consent form before undergoing the aeoNose procedure. Participants with a rheumatologist-confirmed clinical diagnosis of SpA were deemed eligible for inclusion in the patient group and participants without a known inflammatory rheumatic disease for inclusion in the control group. People unable to perform the aeoNose procedure were excluded from the study, along with those who had insufficient command of the Dutch language to fully understand the informed consent. Additionally, it was decided to exclude the subgroup of SpA patients presenting with psoriasis, also known as psoriatic arthritis (PsA), to ensure a more homogenous patient group. To train sufficiently robust prediction models, it was estimated that each group had to comprise at least 50 participants.

Materials

The aeoNose contains three different rigid metal oxide sensors, each of the micro-hotplate type [23]. This design allows modulation of the sensor surface temperature through a 32-step sinusoidal process, ranging between 180° and 340°. Within this range, the sensors function as semiconductors. As the temperature increases, the semiconductor's ability to conduct electricity also increases. Initially, when oxygen molecules adsorb/ ionize at the sensor surface, conductivity remains low. However, the removal of oxygen due to redox reactions with VOCs causes the conductivity to change [26]. These changes can be measured and quantified, resulting in a unique breath profile. Moreover, the modulation of the sensor surface temperature allows for conductivity changes to be obtained at different temperatures [27]. This provides a more accurate measurement of the composition of breath, considering the temperature dependence of the redox reactions. Conductivity is therefore recorded as a function of the temperature modulation.

A small pump within the aeoNose ensures that a steady stream of exhaled air flows over the three sensor surfaces and a Tenax tube, allowing for the measurement of VOCs [27]. A carbon filter in the disposable mouthpiece ensures the inhalation of clean air, while another carbon filter beneath the device ensures that only clean air is passed over the sensors. These measures minimize the risk of cross-infection. Additionally, the aeoNose contains one-way valves, allowing exhalation through the device but preventing any air from returning to the subject [26].

Procedure

All measurements were performed in the same room with the same aeoNose device (AEO18120006) for both patients and controls, following a standardized procedure. The room was only used for aeoNose measurements and to avoid interference, the use of hand-alcohol was prohibited in this room. Participants were instructed to inhale and exhale through the disposable mouthpiece of the aeoNose for five minutes. During this time, the circulation of air should only go through the aeoNose device. Therefore, nose breathing was interfered with by using a nose clamp and patients were asked to enclose the lips over the mouthpiece at all times. Patients were also instructed to hold the aeoNose device in their hands, positioning themselves as preferred. During the measurement, patients were informed when they reached the 2.5-minute time limit, indicating they were halfway done. After 5 minutes of breathing, the end of the measurement was indicated by an audio signal from the aeoNose device. At this point, participants were asked a yes/no question regarding their willingness to perform the diagnostic procedure again. The perceived burden was assessed on a scale from 0 to 10, with 0 indicating no discomfort at all and 10 indicating unbearable discomfort. This information was recorded in Castor Electronic Data Capture, EDC (v2023.4.5.0). After the measurement, the aeoNose was set aside and the sensor surfaces were regenerated with clean air. Then, the internal Tenax tube that gathered VOCs during the measurement was heated, while the VOCs were released [28]. Finally, the data was sent over to the iPad through Bluetooth, after which it was sent to the data center of the eNose company for analysis.

After undergoing the aeoNose procedure, demographic information on age, gender, height, body weight, smoking status, comorbidities, and the last time that the patient had eaten was collected and captured in Castor EDC. Additional information regarding the disease status of SpA patients was gathered from their electronic patient files and also documented in Castor EDC. This information included details about the peripheral or axial nature of the disease (SpA location), the presence of the HLA-B27 gene, CRP levels, the erythrocyte sedimentation rate (ESR), disease activity, and the use of csDMARDs or bDMARDs. The diagnosis, SpA location, and activity of the disease were also verified by the specialist consulted on the day of the measurement, using a short case record form that the physician was asked to fill out (Appendix B).

Statistical analysis

Descriptive statistics

Descriptive statistics were carried out to describe the study population and to explore differences between the study groups. This firstly involved exporting the data that was captured in Castor EDC to a CSV file, which was then imported into RStudio (2023.06.2+561). Subsequently, various tests were conducted to assess whether there were statistically significant differences in the collected variables between both groups. Since the groups are independent, a t-test was performed for all continuous variables that were normally distributed. In cases where a normal distribution could not be assumed, the Mann-Whitney U-test was utilized to compare the groups. For categorical variables, a Chi-squared test was employed. Differences with a p-value below 0.05 were considered significant.

In the patient group, SpA-related details were collected along with the general characteristics. Therefore, a separate table describing disease-specific characteristics was generated. For continuous variables that followed a normal distribution, the mean and standard deviation were computed. When a normal distribution could not be assumed, the median and interquartile range were provided. Categorical variables were presented as counts of subjects and the corresponding percentages.

Generating the prediction models

VOC data preparation and machine learning-based training were performed by the eNose company using the proprietary Aethena software package [23]. During each measurement, a matrix was recorded containing thousands of data points. This data was first standardized, due to the slight variations that exist between sensors among aeoNoses. Subsequently, data was compressed using a Tucker3 algorithm, resulting in vectors representing one of the seven sensor combinations of the three metal oxide sensors in the aeoNose device. Redundant information and noise were removed in the compressed vectors, maintaining the information concerning the distinction between healthy controls and patients.

To create a prediction model, various machine learning (ML) methods can be used. In this study, the eNose company entered the vectors that were created into a supervised random forest (RF) model, while the true disease status was known. This resulted in the generation of multiple possible RF models, each with its specific performance measures. The output of each model included a ranked receiver operating characteristic (ROC) curve, with performance measured by the area under the receiver operating characteristic curve (AUC-ROC), sensitivity, and specificity. A combination of the best models was selected based on ranked AUC-ROCs, resulting in so-called 'judge' models. The selected judge models were validated using 10-fold cross-validation, as the limited sample size did not allow for splitting the sample into a separate training and test set. For further analyses, the algorithm of the best performing model was used. This algorithm produced a prediction value for each study subject from the RF, ranging between -1 and +1, with higher values indicating a higher likelihood of SpA.

Next, three different logistic regression models were constructed in RStudio, all employing the true disease status as the dependent variable. The first model consisted of a univariate logistic regression, incorporating only the aeoNose value as the independent variable. For the second (multivariate) logistic regression model, solely clinical variables were added as independent variables. Initially, all candidate variables were introduced to the model. These included age, gender, body mass index (BMI), smoking status, the presence of comorbidities, and the last moment of food consumption. Subsequently, backward selection was employed to identify variables with significant impact on the model's fit, as determined by the -2log likelihood [29]. Significance for retaining the variables in the model was indicated by a p-value below 0.15. SpA-related features such as location and activity of the disease were not included in the model, as they did not apply to control subjects. Regarding medication use, it could be assumed that no healthy control subject uses either csDMARDs or bDMARDs, while a considerable number of patients do, making these unsuitable as fair predictors. Finally, a third multivariate logistic regression was conducted, incorporating the clinical variables from the second model along with the aeoNose classification value. The classification value was multiplied by 10 before including it in the regression. This adjustment allowed the odds ratio to be interpreted as the odds of having SpA for every 0.1 increment in the classification value, rather than for an increment of 1. This interpretation is more meaningful, given that the aeoNose value can only range from -1 to +1. Backward selection was then employed again to determine which variables should be retained in the model ($p < 0.15$). This process was repeated because the addition of the classification value alters the interactions between the variables in the regression model, potentially leading to different variables becoming significant in predicting disease status. Before conducting the backward selection, it was decided to always incorporate age in the final model, as this variable is readily available and commonly used in similar studies.

For each model, a different cut-off point was chosen, such that healthy controls and subjects were optimally separated [23]. The decision on the optimal threshold is dependent on the purpose of the breath

test. A lower threshold reduces the number of missed SpA cases, while a higher threshold reduces the number of false-positive cases. The selected cut-off point consequently determines the sensitivity and specificity of the model. A feasible aim for the aeoNose is to utilize it as a pre-screening tool, preferably within general practitioners' (GP) offices. Here, patients with suspected SpA can undergo an aeoNose measurement. The results of the measurement can assist the GP in determining whether referral to the rheumatology department for further evaluation is warranted. To realize this in practice, the number of false negatives as measured by the aeoNose should be minimized, to avoid missed diagnoses. Therefore, favoring high sensitivity over specificity is preferred when deciding on a cut-off point, along with a high NPV. On the other hand, excessive referrals of healthy patients are also undesirable. Hence, the number of false positives should be maintained at an acceptable level.

The exact cut-off point for each model in this study was determined in collaboration with a rheumatologist, leveraging their clinical expertise. They stated that the objective for each model should be to achieve a sensitivity of at least 90%, ideally 95%, to be clinically relevant as a pre-screening tool. To compute the ROC-AUC of all models, the sensitivity and specificity for each possible cut-off point were calculated to serve as inputs for the ROC, using the formulas in Figure 1 [30]. In addition, the positive predictive value (PPV) and negative predictive value (NPV) were calculated using the formulas in Figure 1 [31].

		Actual Status		
		Disease Positive	Disease Negative	
Assignment Status	Test Positive	TP	FP	PPV =TP/(TP + FP)
	Test Negative	FN	TN	NPV =TN/(TN + FN)
		Sensitivity =TP/(TP + FN)	Specificity =TN/(TN + FP)	

Figure 1: Formulas used to measure the performance of the various models, with the FN being the false negatives, FP false positives, TN true negatives, and TP true positives [31].

As the RF models generated by the eNose company were validated using 10-fold cross-validation, this technique was also applied in the training process of the various logistic regression models. The performance of these models was evaluated using Cohen's Kappa, as this accuracy statistic accounts for imbalanced study arms [32]. Given the considerable variation in the size of the study groups, Cohen's Kappa was deemed appropriate to use in this study. To finalize the analysis, the AUC-ROCs of all three models were compared using DeLong's test. Differences with a p-value below 0.05 were considered significant.

Results

A total of 320 subjects were eligible for inclusion in this study (Figure 2). Of these subjects, 34 had to be excluded due to failed measurements. Another 36 were excluded because they performed the breath test with a different aeoNose device (AEO19010006), which did not collect sufficient data to be included in the analysis. Finally, 11 subjects were excluded because they were initially classified as healthy but were later found to have a rheumatic disease. As a result, the final study cohort comprised 239 subjects, including 180 healthy controls and 59 SpA patients.

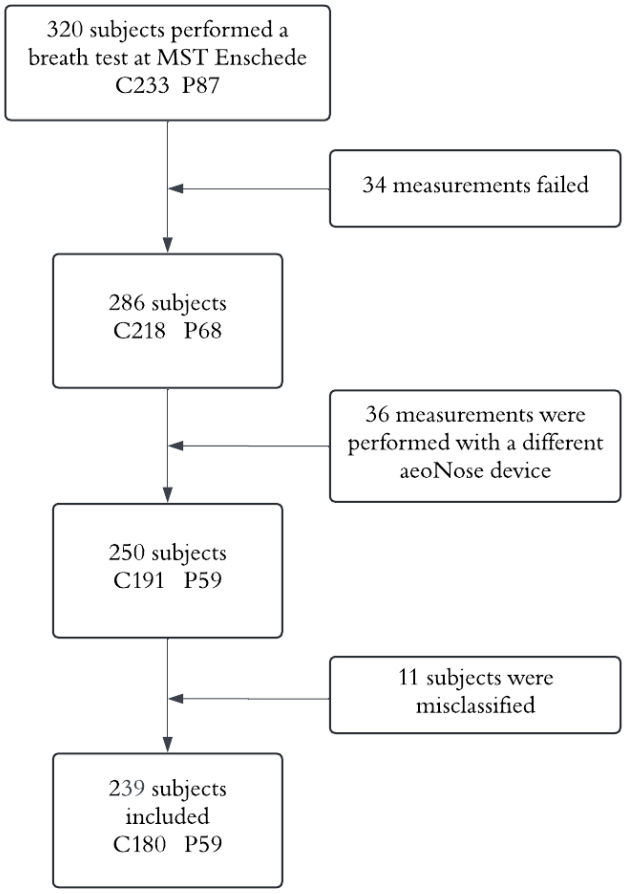

Figure 2: Flowchart of the inclusion process. C = healthy controls, P = SpA patients

Table 1 provides the baseline characteristics of the study samples. Both groups were similar in terms of general characteristics, but SpA patients were significantly more likely to be male (p=0.022). The smoking status of one participant in the control group was missing. Since smoking status was included as an independent variable in the multivariate logistic regression, this participant had to be removed from further analyses.

Concerning the experience of participants, >90% of them were willing to perform the breath test again. However, some complications were observed during the measurements. Firstly, the nose clamp was quite unstable, causing it to fall off regularly. Additionally, the disposable mouthpiece often became briefly disconnected, as it was not securely attached to the device. Despite these minor issues, participants' overall experience was positive. This is reflected in the median discomfort scores, which were 2 for both patients and controls on a scale where 0 indicated no discomfort at all and 10 indicated unbearable discomfort.

Table 1: Baseline characteristics of the study population (N=239)

	SpA patients N=59	**Healthy controls N=180**	**P-value**
Mean age (SD)	50.3 (15.8)	53.3 (17.6)	0.213
Gender N(%)			0.022
Male	36 (61.0)	77 (42.8)	
Female	23 (39.0)	103 (57.2)	
Mean BMI (SD)	26.6 (4.60)	27.0 (5.18)	0.514
Smoking N(%)			0.824
Yes	7 (11.9)	28 (15.6)	
No	29 (49.2)	81 (45.0)	
Stopped	23 (39.0)	70 (38.9)	
Missing	0 (0.0)	1 (0.6)	
Comorbidities N(%)[(1)]			0.930
Yes	36 (61.0)	113 (62.8)	
No	23 (39.0)	67 (37.2)	
Last time eating N(%)			0.878
Less than 3 hours	42 (71.2)	132 (73.3)	
More than 3 hours	17 (28.8)	48 (26.7)	
Willingness to perform again N(%)			0.357
Yes	54 (91.5)	164 (91.1)	
No	3 (5.1)	4 (2.2)	
Missing	2 (3.4)	12 (6.7)	
Median discomfort score (IQR)	2 (1-4)	2 (1-4)	0.744[(2)]

[(1)]Specified in Appendix C.

[(2)]The p-value was determined through a Mann-Whitney U test instead of a t-test.

The clinical characteristics of the SpA patients are displayed in Table 2. Among the overall patient group, 41 individuals were identified with axial SpA (69.5%), while only 15 patients had peripheral SpA (25.4%). Additionally, the majority of patients tested positive for the HLA-B27 gene (66.1%).

In total, 34 patients received SpA-specific medications (57.6%). Among them, 18 received bDMARDs (30.5%), especially TNF inhibitors (27.1%), while the remaining 16 patients used csDMARDs (27.1%). The

median CRP level was determined to be 2, with an interquartile range of 1 to 6, indicating low disease activity.

Table 2: Characteristics of SpA patients (N=59)

	SpA patients
SpA location N(%)	
Axial	41 (69.5)
Peripheral	15 (25.4)
Missing	3 (5.1)
HLA-B27 N(%)	
Yes	39 (66.1)
No	14 (23.7)
Missing	6 (10.2)
Median ESR (IQR)	6.5 (2-12.8)
Median CRP (IQR)	2 (1-6)
Active illness N(%)	
Yes	10 (16.9)
No	35 (59.3)
Missing	14 (23.7)
Medication use N(%)	
csDMARD	16 (27.1)
Methotrexate	7 (11.9)
Sulfasalazine	5 (8.5)
Hydroxychloroquine	1 (1.7)
Prednisone	3 (5.1)
bDMARD	18 (30.5)
TNFi inhibitor	16 (27.1)
Abatacept	1 (1.7)
IL6-inhibitor	1 (1.7)
Other	52 (88.1)

Prediction models

In the ML prediction modeling analysis, 180 healthy controls and 59 SpA patients were included. Due to the exclusion of one participant in the control group with a missing smoking status, the subsequent logistic regression models comprised 179 healthy controls and 59 patients.

The best RF model of the eNose company assigned a value between -1 and 1 to each participant, indicative of their disease status. These predicted values for all participants are displayed in the scatterplot in Figure 3. In this plot, red squares represent SpA patients, while green squares represent healthy controls. The dashed line indicates the optimal cut-off point that was used by the eNose company, set at a value of -0.13. At this threshold, the model achieved a sensitivity of 92%, a specificity of 88%, a PPV of 71%, and an NPV of 97%. The AUC-ROC was 0.95 (95% CI 0.92-0.99). Participants with classification values below the threshold are predicted to be healthy, while those above the threshold are predicted to be patients.

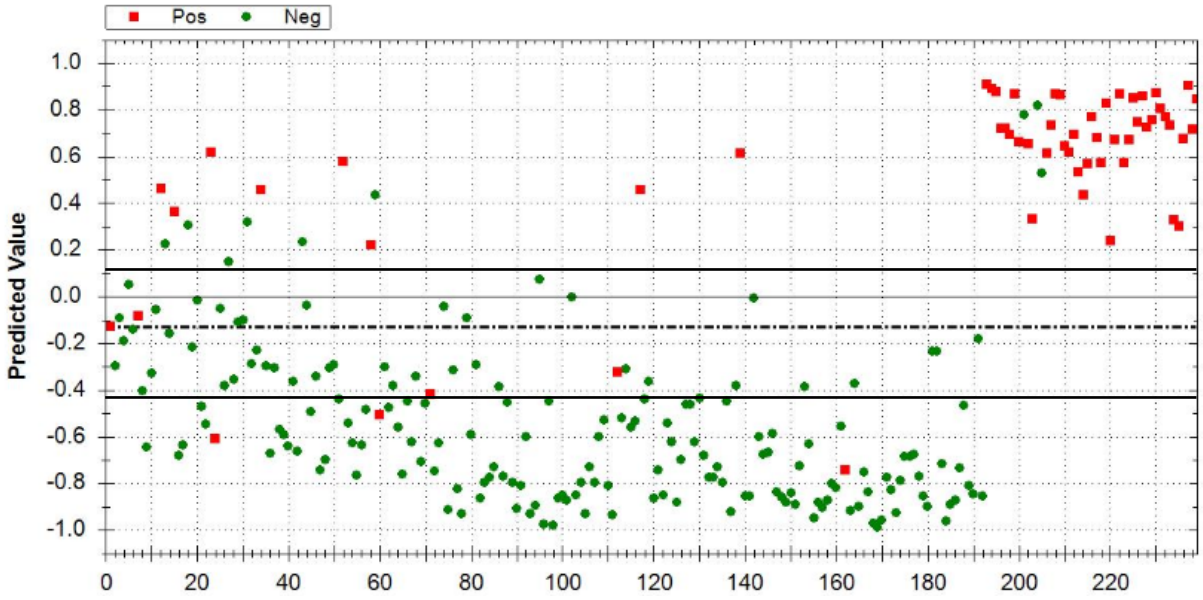

Figure 3: Scatterplot of the predicted value per participant. The dashed line represents the threshold determined by the eNose company. Solid lines represent the thresholds used in the univariate logistic regression.

To illustrate the effects of changing the threshold, the eNose prediction value was entered into the first univariate logistic regression. To achieve a higher specificity and PPV, the threshold was increased to 0.11 (Figure 3), which corresponds to a probability ≥38% of SpA. With this cut-off point, the sensitivity reached 88%, the specificity 95%, the PPV 85%, and the NPV 96%. Since this study aims to investigate the potential of the aeoNose as a pre-screening tool, the definite model should exhibit high sensitivity and NPV. To meet these criteria, the threshold was finally set to -0.42 (Figure 3), corresponding to a probability ≥5% of SpA. This adjustment resulted in a sensitivity of 95%, a specificity of 70%, a PPV of 51%, and an NPV of 98%. Since this first logistic regression model included only the aeoNose classification value, the AUC-ROC remained 0.95 (95% CI 0.92-0.99).

The second logistic regression model included only relevant clinical variables, determined by backward selection. Table 3 shows that age and gender significantly predicted the presence of SpA. Including these variables in the model resulted in a sensitivity of 95%, a specificity of 9%, a PPV of 26%, and an NPV of 85%. The corresponding AUC was 0.63 (95% CI 0.54-0.71). A probability of SpA ≥14% was considered a positive test result in this case.

The third model combined both eNose values and relevant clinical variables, again determined through backward selection. Table 3 shows that the aeoNose breath value is strongly associated with the presence of SpA (OR 1.62, 95% CI 1.46-1.86). Additionally, male participants have a 3-fold higher odds of having SpA (OR 3.02, 95% CI 1.02-9.73). Age was not significantly associated with the presence of SpA in this model (p = 0.47), but was still incorporated in the final model as described in the method section. Incorporating these variables in the regression resulted in an only marginally higher AUC of 0.96 (95% CI 0.93-0.99). The sensitivity of the model was 95%, the specificity 73%, the PPV 53%, and the NPV 98%. For this model, a probability of SpA ≥5% was considered a positive test result, corresponding to a threshold of -0.53.

Table 3: Results of the backward selection for model 2 and model 3

Variable	Model 2	p-value	Model 3	p-value
aeoNose value[(1)]	-	-	1.62 (1.46-1.86)	<0.001
Age	0.99 (0.97-1.00)	0.106	1.01 (0.98-1.05)	0.47
Gender (male)	2.28 (1.24-4.27)	0.0089	3.02 (1.02-9.73)	0.052

Data are presented as odds ratio (95% confidence interval), along with the p-value. Constant model 2: -0.79, model 3: -2.26. Cut-off point: 0.15. [(1)]The aeoNose value was multiplied by 10 to obtain the odds per 0.1 increment.

Figure 4 shows the confusion matrices of classifications made by each model, which were used to calculate the sensitivity, specificity, PPV, and NPV. Here, the prediction for every participant is depicted against their true disease status. Table 4 summarizes all relevant outcome measures, along with Cohen's Kappa. This measure was used to evaluate the validity of the models, resulting from the 10-fold cross-validation. Cohen's Kappa can range from -1 to +1, with +1 indicating perfect agreement between predictors and outcome and 0 representing the level of agreement expected by random chance [34]. Values below 0 are also possible, but unlikely to occur in practice. The reported Cohen's Kappa is the average result across all 10 folds of the validation.

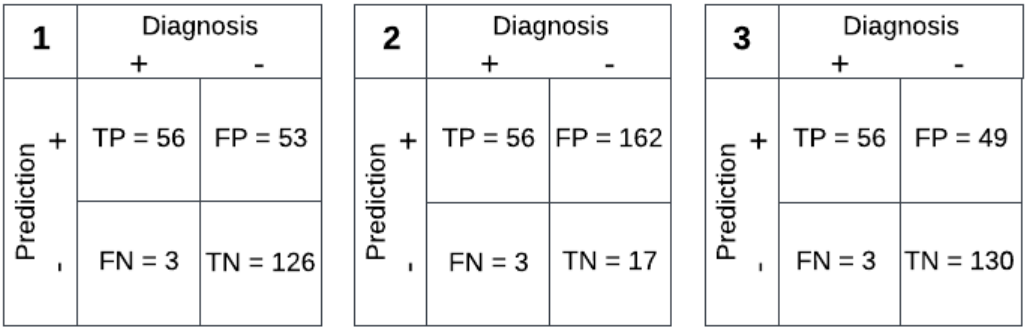

Figure 4: Outcomes of the three logistic regression models (left upper corner). Model 1 contains only the aeoNose classification value, model 2 solely clinical variables, and model 3 a combination of both clinical variables and the aeoNose value. TP =true positives, FP = false positives, FN = false negatives, TN = true negatives.

Table 4: Performance of the three investigated logistic regression models

Model	Cut-off	Sensitivity	Specificity	PPV	NPV	AUC-ROC (95% CI)	Kappa
1	-0.42	95%	70%	51%	98%	0.95 (0.92-0.99)	0.82
2	0.11	95%	9%	26%	85%	0.63 (0.54-0.71)	0
3	-0.53	95%	74%	55%	98%	0.96 (0.93-0.99)	0.84

Model 1 contains only the aeoNose classification value, model 2 solely clinical variables, and model 3 a combination of both clinical variables and the aeoNose value. PPV: positive predictive value, NPV: negative predictive value, AUC-ROC: area under the receiver operating curve, Kappa: mean value of the 10 folds, indicative of the model's fit.

To visualize the performance of the various models, Figure 5 presents the ROC curves of the three models. In this figure, the sensitivity is plotted on the y-axis against the 1 - specificity on the x-axis.

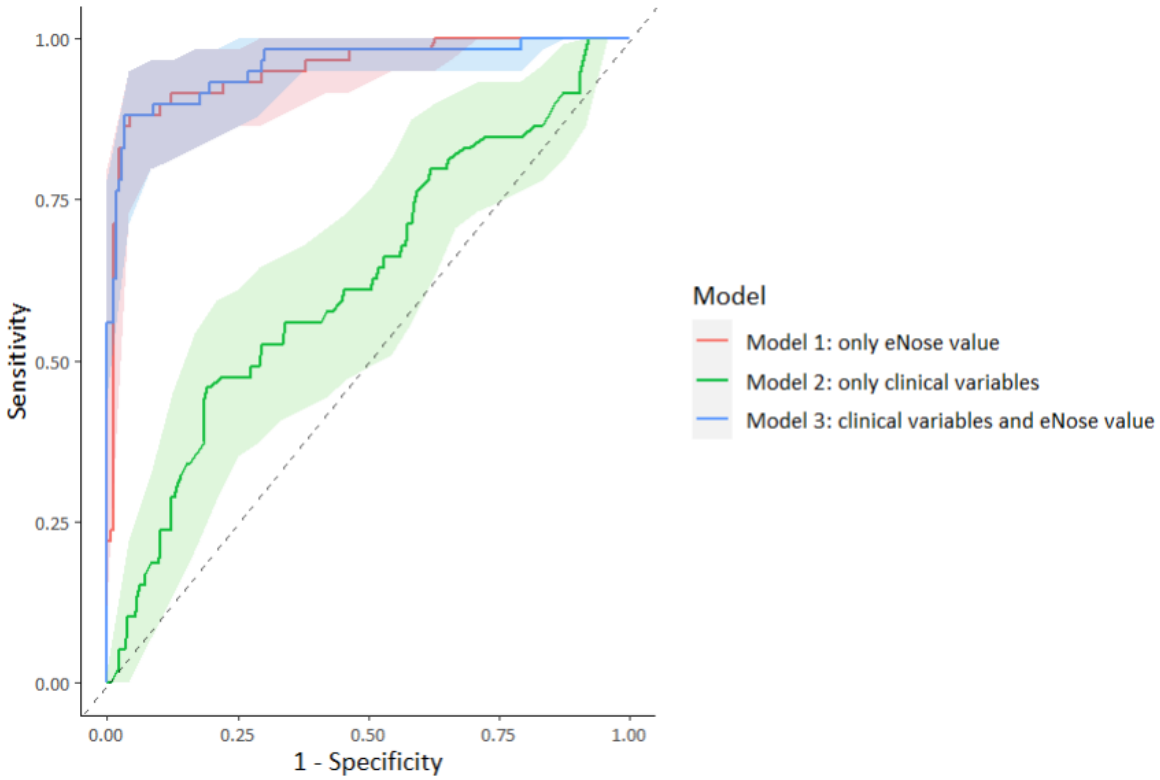

Figure 5: Combined ROC curve for models 1, 2, and 3. Shades represent 95% CI's.

To conclude the analysis, several tests were conducted to determine whether the overall AUC-ROCs of the various models significantly differed from each other. The AUC of the aeoNose values-only model was significantly higher than that of the clinical variables-only model ($p<0.001$). Comparing the aeoNose values-only model with the combined model revealed no significant difference in AUC ($p=0.47$). Finally, the AUC of the clinical variables-only model was significantly lower than that of the combined model ($p<0.001$).

Discussion

This pilot study investigated the ability of the aeoNose to discriminate between SpA patients and healthy controls through exhaled breath analysis. With the prediction model that contained solely aeoNose classification values, a sensitivity of 95% and an NPV of 98% were found in the training sample. Diagnostic accuracy, expressed as AUC-ROC, was found to be 0.95. Adding readily available clinical parameters did not change the sensitivity and NPV but slightly reduced the number of false positives, increasing the AUC-ROC to 0.96. However, as the performance of the initial aeoNose model was already very high, this increase was not statistically significant (p=0.47). These results indicate that the aeoNose was highly effective in distinguishing the breath prints of healthy controls from those of SpA patients. However, these findings need to be validated in larger and independent samples.

The sensitivity, specificity, NPV, and PPV observed in the various models were determined by the chosen cut-off point, as this influences the balance between the number of false-positive cases versus the number of missed cases concerning SpA. The decision on these thresholds is subjective and dependent on the aim of the test. When the consequences of late disease detection are severe, it is crucial to minimize the number of missed cases and focus should therefore be on obtaining a high sensitivity and NPV. Conversely, when used in a diagnostic setting, the number of false-positive test results should be minimized. In this case, the focus should be on achieving a high specificity and PPV.

Since previous studies indicated that the aeoNose might not perform well enough as a standalone diagnostic tool [22, 35], the current study focused on its use as a pre-screening tool in a point-of-care setting. This approach to the use of an eNose aims to reduce diagnostic delays in SpA while at the same time preventing too many unnecessary referrals of healthy individuals to hospitals, thereby alleviating pressure on the waiting lists and specialists. However, the excellent predictive results obtained in the current study indicate that the aeoNose may even have potential for use in a diagnostic setting. This could be particularly effective for SpA, a disease that is challenging to diagnose in its early stages. Existing diagnostic techniques for SpA include HLA-B27 status, the ASAS criteria, and MRI, which exhibit sensitivities and specificities around 43% and 91% [36], 83% and 84% [7], and 79% and 89% [37, 38], respectively. Given the sensitivity of 88% and the specificity of 95% found in the current training sample, the aeoNose seems to outperform the existing techniques. Employing the aeoNose as a diagnostic tool in a hospital setting would be beneficial, as it would eliminate the need to repeat the calibration process at every external location separately.

The diagnostic value of the eNose has already been evaluated across various conditions, including respiratory diseases, infectious diseases, and cancer. However, studies on using an eNose for diagnosing inflammatory diseases are limited [39]. Brekelmans et al. (2026) investigated whether an eNose could differentiate between patients with rheumatoid arthritis (RA), PsA, and healthy controls. Their study revealed that the breath prints of RA and PsA patients could be differentiated from those of healthy controls with accuracies of 71% and 69% in the training set, respectively. The current study marks the first exploration of the aeoNose's ability to discriminate healthy controls from SpA patients.

Strengths

An important strength of this study is the comparable groups in terms of general characteristics. With the exception of SpA patients being predominantly male, the groups showed no statistically significant differences in age, BMI, smoking status, the presence of comorbidities, and the last moment of food consumption. Therefore, the exhaled breath profiles were not likely to be substantially influenced by these

factors, making the findings more robust. Medication usage could also be a confounding factor, as they are typically used among most patients. It is possible that the aeoNose may not detect variations in breath profiles due to inflammation, but instead also measures the presence of SpA-specific medications. However, in this study, 34 patients were using SpA-specific medications out of the total patient group of 59. Considering that nearly half of the patients were not on SpA medication, the potential confounding effect is also expected to be limited. The remaining characteristics of the patient group in this study are also representative of the general SpA population, as seen by rheumatologists in the Netherlands [33].

Another strength of this study is the systematic and standardized approach that was used to complete the measurements. All measurements were performed in the same room under stable, controlled conditions. The researcher overseeing the measurements ensured the correct use of the nose clamp and verified that participants were breathing through the mouthpiece, without air escaping along the sides. If accurate measurement could not be guaranteed, such as when too many breaks were taken, the participant was excluded from the study.

Limitations

This study provides only the first step in exploring the ability of the aeoNose to discriminate between healthy controls and SpA patients. The RF model and the logistic regression models used in this study were only validated using 10-fold cross-validation, meaning that no independent test sample was used to validate the outcomes. The main consideration for not splitting the data into a training and testing set separately was the fact that the patient group comprised only 59 individuals. Splitting the data would therefore result in the inclusion of only a few patients in de dataset, most likely leading to inaccurate validation results [40]. While 10-fold cross-validation prevents overfitting of the models to some extent, it cannot be guaranteed that the models are completely free of overfitting. Therefore, externally validating the results of the performance of the algorithm in a future study is crucial.

When comparing the outcomes of this study to similar eNose studies, it is evident that the initial AUC-ROC value of 0.95 found here is exceptionally high. Other studies typically reported AUC-ROCs ranging between 0.70 and 0.85 [22, 28, 39]. The discrepancy in results can partly be explained by a flaw in the study design. The control group was primarily collected during an earlier study by D. Gerritsen et al. (2023), with measurements primarily executed between January 2022 and March 2023. Data collection for the current SpA patient group mostly occurred between February and April 2024. Therefore, most healthy controls were included in the study significantly earlier than SpA patients. The separation plot (Figure 3) illustrates this time discrepancy. Healthy controls are primarily located on the left side of the plot, whereas patients are mostly located in the right upper corner. This indicates the presence of distinct clusters of data points. However, for accurate training of a prediction model, it is important to measure both positive and negative samples simultaneously in time [41]. When data collection occurs at different times for various groups, the model might learn from temporal patterns in the data instead of actual breath data patterns, leading to inaccurate performance. This effect is also illustrated in Figure 3, where the few controls measured in the last period of the study are all assigned high eNose values, along with the patients measured during that time. Because this period predominantly included patients, the algorithm assumes everyone measured during this time is a patient. Due to this uneven distribution of data, the results of the current study are likely to overestimate the diagnostic accuracy of the aeoNose in independent samples. Simultaneously measuring patients and controls would likely have led to lower performance of the prediction model compared to its current performance, as it would eliminate the confounding effect of data clusters and ensure the model is trained on solely breath data.

Another limitation of this study is the limited inclusion of clinical variables in the multivariate logistic regression. While laboratory findings like CRP levels and ESR were accessible for SpA patients, healthy controls were not subjected to blood tests. Consequently, this information could not be included in the logistic regression, leaving only general characteristics to be considered. Among these general characteristics, smoking behavior was categorized in the analysis as current, past, or never smoker. However, the preferred method to capture smoking behavior is through pack years. In this measure, both the quantity and duration of smoking are considered, providing a more accurate reflection of its impact. This is important as smoking affects SpA in various ways [42]. Pack years could not be utilized in this study, as the amount smoked by past smokers was not recorded. This would have resulted in a pack year value of 0 for all past smokers, inaccurately representing their smoking history.

Future research

Further research is recommended to gain more insight into the discriminatory abilities of the aeoNose. While this pilot has demonstrated the ability of the aeoNose to differentiate between breath prints of healthy controls and those of SpA patients in the training set, it is important to acknowledge the described limitations that are likely to have affected the robustness of findings. Therefore, validation of the results in a new study is necessary. It would be advised to include a larger sample size in such a study and to also conduct subgroup analyses, for example between SpA patients with active disease and those without, to investigate the influence on the performance of the aeoNose. Moreover, the collection of blood work from both patients and healthy controls would enable consideration of more clinical variables in the multivariate logistic regression, potentially leading to improved performance.

Lastly, it cannot be conclusively stated that the diagnostic performance of the aeoNose is solely attributed to the pathological processes represented by endogenous VOCs. There is a possibility that the aeoNose captures other factors characteristic of SpA, enabling the distinction between patients and healthy controls. Validating the device in a different setting with a different population would address these assumptions, allowing more robust conclusions.

Conclusion

Based on this study it can be concluded that the aeoNose shows promising ability to discriminate between SpA patients and healthy controls and could possibly be used as a diagnostic tool in the future. Diagnostic accuracy can be slightly enhanced by incorporating age and gender into a multivariate logistic regression model, rather than including solely aeoNose classification values based on exhaled breath data. However, further research is needed to validate these findings.

References

1. Carron P, De Craemer AS, Van den Bosch F. Peripheral spondyloarthritis: a neglected entity-state of the art. RMD Open. 2020;6(1). doi: 10.1136/rmdopen-2019-001136.
2. Wendling D, Claudepierre P, Prati C. Early diagnosis and management are crucial in spondyloarthritis. Joint Bone Spine. 2013;80(6):582-5. doi: 10.1016/j.jbspin.2013.03.003.
3. Molto A, Sieper J. Peripheral spondyloarthritis: Concept, diagnosis and treatment. Best Pract Res Clin Rheumatol. 2018;32(3):357-68. doi: 10.1016/j.berh.2019.02.010.
4. Torgutalp M, Poddubnyy D. Emerging treatment options for spondyloarthritis. Best Pract Res Clin Rheumatol. 2018;32(3):472-84. doi: 10.1016/j.berh.2019.01.014.
5. Raychaudhuri SP, Deodhar A. The classification and diagnostic criteria of ankylosing spondylitis. J Autoimmun. 2014;48-49:128-33. doi: 10.1016/j.jaut.2014.01.015.
6. Carvalho PD, Machado PM. How to investigate: Early axial spondyloarthritis. Best Pract Res Clin Rheumatol. 2019;33(4):101427. doi: 10.1016/j.berh.2019.07.001.
7. Rudwaleit M, van der Heijde D, Landewé R, Listing J, Akkoc N, Brandt J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis. 2009;68(6):777-83. doi: 10.1136/ard.2009.108233.
8. Redeker I, Siegmund B, Ghoreschi K, Pleyer U, Callhoff J, Hoffmann F, et al. The impact of extra-musculoskeletal manifestations on disease activity, functional status, and treatment patterns in patients with axial spondyloarthritis: results from a nationwide population-based study. Ther Adv Musculoskelet Dis. 2020;12:1759720x20972610. doi: 10.1177/1759720x20972610.
9. Chen B, Li J, He C, Li D, Tong W, Zou Y, et al. Role of HLA-B27 in the pathogenesis of ankylosing spondylitis (Review). Mol Med Rep. 2017;15(4):1943-51. doi: 10.3892/mmr.2017.6248.
10. Sharip A, Kunz J. Understanding the Pathogenesis of Spondyloarthritis. Biomolecules. 2020;10(10). doi: 10.3390/biom10101461.
11. Stolwijk C, van Onna M, Boonen A, van Tubergen A. Global Prevalence of Spondyloarthritis: A Systematic Review and Meta-Regression Analysis. Arthritis Care Res (Hoboken). 2016;68(9):1320-31. doi: 10.1002/acr.22831.
12. Walsh JA, Magrey M. Clinical Manifestations and Diagnosis of Axial Spondyloarthritis. J Clin Rheumatol. 2021;27(8):e547-e60. doi: 10.1097/rhu.0000000000001575.
13. Navarro-Compán V, Sepriano A, El-Zorkany B, van der Heijde D. Axial spondyloarthritis. Ann Rheum Dis. 2021;80(12):1511-21. doi: 10.1136/annrheumdis-2021-221035.
14. Poddubnyy D. Classification vs diagnostic criteria: the challenge of diagnosing axial spondyloarthritis. Rheumatology (Oxford). 2020;59(Suppl4):iv6-iv17. doi: 10.1093/rheumatology/keaa250.
15. Bittar M, Khan MA, Magrey M. Axial Spondyloarthritis and Diagnostic Challenges: Over-diagnosis, Misdiagnosis, and Under-diagnosis. Curr Rheumatol Rep. 2023;25(3):47-55. doi: 10.1007/s11926-022-01096-0.
16. van Gaalen FA, Rudwaleit M. Challenges in the diagnosis of axial spondyloarthritis. Best Pract Res Clin Rheumatol. 2023;37(3):101871. doi: 10.1016/j.berh.2023.101871.
17. Zhao SS, Pittam B, Harrison NL, Ahmed AE, Goodson NJ, Hughes DM. Diagnostic delay in axial spondyloarthritis: a systematic review and meta-analysis. Rheumatology (Oxford). 2021;60(4):1620-8. doi: 10.1093/rheumatology/keaa807.
18. Hay CA, Packham J, Ryan S, Mallen CD, Chatzixenitidis A, Prior JA. Diagnostic delay in axial spondyloarthritis: a systematic review. Clin Rheumatol. 2022;41(7):1939-50. doi: 10.1007/s10067-022-06100-7.
19. Boonen A, Chorus A, Miedema H, van der Heijde D, Landewé R, Schouten H, et al. Withdrawal from labour force due to work disability in patients with ankylosing spondylitis. Ann Rheum Dis. 2001;60(11):1033-9. doi: 10.1136/ard.60.11.1033.

20. Boonen A, Brinkhuizen T, Landewé R, van der Heijde D, Severens JL. Impact of ankylosing spondylitis on sick leave, presenteeism and unpaid productivity, and estimation of the societal cost. Ann Rheum Dis. 2010;69(6):1123-8. doi: 10.1136/ard.2009.116764.
21. Aouad K, El-Zorkany B. Treat-to-Target in Axial Spondyloarthritis: Are we there yet? Mediterr J Rheumatol. 2022;33(Suppl 1):137-41. doi: 10.31138/mjr.33.1.137.
22. Brekelmans MP, Fens N, Brinkman P, Bos LD, Sterk PJ, Tak PP, et al. Smelling the Diagnosis: The Electronic Nose as Diagnostic Tool in Inflammatory Arthritis. A Case-Reference Study. PLoS One. 2016;11(3):e0151715. doi: 10.1371/journal.pone.0151715.
23. Kort S, Brusse-Keizer M, Gerritsen JW, van der Palen J. Data analysis of electronic nose technology in lung cancer: generating prediction models by means of Aethena. J Breath Res. 2017;11(2):026006. doi: 10.1088/1752-7163/aa6b08.
24. Kort S, Brusse-Keizer M, Schouwink H, Citgez E, de Jongh FH, van Putten JWG, et al. Diagnosing Non-Small Cell Lung Cancer by Exhaled Breath Profiling Using an Electronic Nose: A Multicenter Validation Study. Chest. 2023;163(3):697-706. doi: 10.1016/j.chest.2022.09.042.
25. van de Goor R, van Hooren M, Dingemans AM, Kremer B, Kross K. Training and Validating a Portable Electronic Nose for Lung Cancer Screening. J Thorac Oncol. 2018;13(5):676-81. doi: 10.1016/j.jtho.2018.01.024.
26. Enose company. User manual. Zupthen 2020.
27. Saktiawati AMI, Stienstra Y, Subronto YW, Rintiswati N, Sumardi, Gerritsen JW, et al. Sensitivity and specificity of an electronic nose in diagnosing pulmonary tuberculosis among patients with suspected tuberculosis. PLoS One. 2019;14(6):e0217963. doi: 10.1371/journal.pone.0217963.
28. Kort S. Detection of lung cancer in exhaled breath with electronic nose technology Enschede University of Twente 2022.
29. Kort S, Brusse-Keizer M, Gerritsen JW, Schouwink H, Citgez E, de Jongh F, et al. Improving lung cancer diagnosis by combining exhaled-breath data and clinical parameters. ERJ Open Res. 2020;6(1). doi: 10.1183/23120541.00221-2019.
30. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol. 2010;5(9):1315-6. doi: 10.1097/JTO.0b013e3181ec173d.
31. Monaghan TF, Rahman SN, Agudelo CW, Wein AJ, Lazar JM, Everaert K, et al. Foundational Statistical Principles in Medical Research: Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value. Medicina (Kaunas). 2021;57(5). doi: 10.3390/medicina57050503.
32. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.
33. Stolwijk C, Boonen A, van Tubergen A, Reveille JD. Epidemiology of spondyloarthritis. Rheum Dis Clin North Am. 2012;38(3):441-76. doi: 10.1016/j.rdc.2012.09.003.
34. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb). 2012;22(3):276-82.
35. Hulshof IM. Ability of the electronic nose to discriminate between patients with rheumatoid arthritis and controls. 2021.
36. García-Salinas R, Ruta S, Chichande JT, Magri S. The Role of HLA-B27 in Argentinian Axial Spondyloarthritis Patients. J Clin Rheumatol. 2022;28(2):e619-e22. doi: 10.1097/rhu.0000000000001763.
37. Jones A, Bray TJP, Mandl P, Hall-Craggs MA, Marzo-Ortega H, Machado PM. Performance of magnetic resonance imaging in the diagnosis of axial spondyloarthritis: a systematic literature review. Rheumatology (Oxford). 2019;58(11):1955-65. doi: 10.1093/rheumatology/kez172.
38. Pedersen SJ, Weber U, Ostergaard M. The diagnostic utility of MRI in spondyloarthritis. Best Pract Res Clin Rheumatol. 2012;26(6):751-66. doi: 10.1016/j.berh.2012.10.005.
39. Farraia MV, Cavaleiro Rufo J, Paciência I, Mendes F, Delgado L, Moreira A. The electronic nose technology in clinical diagnosis: A systematic review. Porto Biomed J. 2019;4(4):e42. doi: 10.1097/j.pbj.0000000000000042.

40. Vabalas A, Gowen E, Poliakoff E, Casson AJ. Machine learning algorithm validation with a limited sample size. PLoS One. 2019;14(11):e0224365. doi: 10.1371/journal.pone.0224365.
41. Yuan W, Beaulieu-Jones BK, Yu KH, Lipnick SL, Palmer N, Loscalzo J, et al. Temporal bias in case-control design: preventing reliable predictions of the future. Nat Commun. 2021;12(1):1107. doi: 10.1038/s41467-021-21390-2.
42. Wendling D, Prati C. Smoking and spondyloarthritis: a bad connection. Rheumatol Int. 2015;35(12):1951-3. doi: 10.1007/s00296-015-3368-0.

Appendix A: informed consent form

Medisch Spectrum Twente
een santeon ziekenhuis

Het opsporen van inflammatoire reumatische ziekten met een elektronische neus.

Postadres
Postbus 50 000
7500 KA Enschede

Bezoekadres
Koningsplein 1
7512 KZ Enschede
T (053) 487 20 00
www.mst.nl

Medisch Spectrum △ Twente
Goedgekeurd
Datum 20-01-20
Hemd-tekening

INFORMATIE OVER DEELNAME AAN EEN WETENSCHAPPELIJK ONDERZOEK

TITEL: Het opsporen van inflammatoire reumatische ziekten met een elektronische neus.
ONDERZOEKERS: dr. H.E. Vonkeman, mw. dr. M. Ghiti Moghadam
CENTRUM: Medisch Spectrum Twente

Inleiding
U wordt gevraagd om deel te nemen aan een medisch-wetenschappelijk onderzoek (zie titel). In dit onderzoek wordt onderzocht of stoffen in de uitgeademde lucht kunnen helpen bij het stellen van de diagnose.
U beslist zelf of u mee wilt doen. Voordat u een beslissing neemt, is het belangrijk dat u over de benodigde informatie beschikt om te kunnen beslissen of u wilt deelnemen. Een arts of één van de onderzoekers zal het onderzoek met u bespreken en al uw vragen beantwoorden. U mag ook met familie en vrienden over uw beslissing praten. Neem alstublieft voldoende tijd om te beslissen.
Dit onderzoek zal uw aandoening niet verbeteren, maar andere patiënten kunnen mogelijk in de toekomst voordeel halen uit de informatie die in dit onderzoek wordt verzameld. Hebt u na het lezen van deze informatie nog vragen? Dan kunt u terecht bij de onderzoeker. Op bladzijde 2 vindt u de contactgegevens.

Doel van het onderzoek
Het doel van dit onderzoek is om te onderzoeken of door met een elektronische neus te ruiken aan uitademingslucht op een minder belastende en snellere manier inflammatoire reumatische ziekten aangetoond of uitgesloten kan worden. Deze methode wordt vergeleken met de huidige onderzoeksmethodes.

Onderzoeksprocedure
Als u deelneemt aan dit wetenschappelijk onderzoek wordt bij u naast de geplande onderzoeken een ademanalyse test gedaan, die niet belastend is.
Op dezelfde dag dat bij u op de polikliniek komt, zal u gevraagd worden om maximaal 5 minuten op uw gewone tempo via de mond in een apparaat te ademen, een elektronische neus. De meting vindt plaats bij het afnamelaboratorium in het ziekenhuis. U hoeft dus geen extra bezoek te brengen aan het ziekenhuis.

Voor- en nadelen van deelname

U heeft zelf geen voordeel van meedoen aan dit onderzoek.
Het onderzoek bestaat enkel uit het opvangen van uitgeademde lucht via een niet belastend apparaat, er worden geen medicijnen toegediend. Er zijn dan ook geen risico's verbonden aan deelname aan dit onderzoek.

Vertrouwelijkheid

Naast bovengenoemde metingen zullen wij mogelijk een aantal andere gegevens van u gebruiken: leeftijd, geslacht, of u rookt of niet, de eventuele aanwezigheid van andere ziekten die invloed kunnen hebben op uw adem, laboratoriumuitslagen, medicijngebruik en de diagnose. Uw persoonlijke gegevens worden enkel in gecodeerde vorm (om ervoor te zorgen dat uw identiteit vertrouwelijk blijft) opgeslagen. Volgens wettelijke bepalingen zullen uw gegevens 15 jaar bewaard worden.

Vrijwillige deelname

Uw deelname aan dit onderzoek is geheel vrijwillig. Als u besluit niet mee te doen, hoeft u verder niets te doen. Indien u wel mee wilt doen, kunt u op elk moment stoppen zonder dat dit gevolgen heeft voor uw verdere behandeling.

Toestemming

Hierbij verklaar ik dat men mij uitgebreid uitleg gegeven heeft over dit onderzoek. Ik heb dit toestemmingsformulier gelezen en begrepen. Al mijn vragen zijn naar tevredenheid beantwoord. Ik stem vrijwillig in met deelname aan dit onderzoek en verzameling van relevante medische informatie. Ik weet dat ik te allen tijde mag stoppen met dit onderzoek en dat dit geen verdere gevolgen voor mij heeft. Ik ontvang een getekend en gedateerd exemplaar van dit toestemmingsformulier.

Ik wil meedoen aan dit onderzoek.

Naam proefpersoon:

Handtekening: Datum : __ / __ / __

Ik verklaar hierbij dat ik deze proefpersoon volledig heb geïnformeerd over het genoemde onderzoek.

Als er tijdens het onderzoek informatie bekend wordt die de toestemming van de proefpersoon zou kunnen beïnvloeden, dan breng ik hem/haar daarvan tijdig op de hoogte.

Naam onderzoeker (of diens vertegenwoordiger):

Handtekening: Datum: __ / __ / __

Bijlage A: Contactgegevens voor Medisch Spectrum Twente

Onderzoekscentrum: Medisch Spectrum Twente

Koningsplein 1

7512 KZ Enschede

Hoofdonderzoekers:

dr. H.E. Vonkeman en mw. dr. M. Ghiti Moghadam.

Bereikbaar tijdens kantooruren op 053-4872450. Voor spoedgevallen buiten kantooruren kunt u contact opnemen met het algemene telefoonnummer 053-4872000 en vragen naar de dienstdoende reumatoloog.

Persoonsgegevens:

Voor meer informatie over de naleving van uw rechten bij de verwerking van uw persoonsgegevens kunt u contact opnemen met de Functionaris voor de Gegevensbescherming van MST, mw. P. van Paridon, telefoon: 06-31751387.

Medisch Spectrum ΔTwente	
Goedgekeurd	
Datum	20-01-20
Hand-tekening	

Appendix B: additional data from the specialist

Aanvullende gegevens eNose
(in te vullen door behandelend reumatoloog):

Patiëntnummer:
Datum meting:
Behandelaar:
Studienummer:

Diagnose	Definitieve diagnose	Actieve ziekte	Eigenschappen
☐ RA	☐ Ja ☐ Nee	☐ Ja ☐ Nee	☐ Erosief ☐ 2-10 grote gezwollen gewrichten ☐ 1-3 kleine gezwollen gewrichten ☐ 4-10 kleine gezwollen gewrichten ☐ >10 gezwollen gewrichten ☐ Ziekteduur ≥6 weken
☐ SpA (ex. PsA)	☐ Ja ☐ Nee	☐ Ja ☐ Nee	☐ Axiaal ☐ Perifeer
☐ PsA	☐ Ja ☐ Nee	☐ Ja ☐ Nee	☐ Erosief
☐ Jicht	☐ Ja ☐ Nee	☐ Ja ☐ Nee	☐ Topheus
☐ PMR	☐ Ja ☐ Nee	☐ Ja ☐ Nee	
☐ SLE	☐ Ja ☐ Nee	☐ Ja ☐ Nee	
☐ SSc	☐ Ja ☐ Nee	☐ Ja ☐ Nee	
☐ Sjögren	☐ Ja ☐ Nee	☐ Ja ☐ Nee	
☐ Inflammatoire artrose	☐ Ja ☐ Nee	☐ Ja ☐ Nee	
☐ AT	☐ Ja ☐ Nee	☐ Ja ☐ Nee	☐ Grote vaten vasculitis
☐ Nieuw consult	Diagnose: ..... Definitief: ☐ Ja ☐ Nee	☐ Ja ☐ Nee	☐ Inflammatoire reumatische aandoening
☐ Anders, namelijk:	☐ Ja ☐ Nee	☐ Ja ☐ Nee	

Opmerkingen:

Appendix C: additional information on comorbidities in the study cohort

	SpA patients N= 59	Healthy controls N= 180
Comorbidities N(%)		
Hypertension	7 (11.9%)	4 (2.2)
Diabetes	2 (3.4)	2 (1.1)
Obesity	0	0
Cardiovascular diseases	7 (11.9)	41 (22.8)
Lung diseases	4 (6.8)	12 (6.7)
Other inflammatory diseases	11 (18.6)	9 (5.0)
Cancer	1 (1.7)	1 (0.6)
Osteoporosis	2 (3.4)	4 (2.2)
Other	26 (44.1)	113 (62.8)